



IMPERIAL INSTITUTE  
OF  
AGRICULTURAL RESEARCH, PUSA.







**PROCEEDINGS**  
**OF THE**  
**ROYAL SOCIETY OF LONDON**

**SERIES A**

**CONTAINING PAPERS OF A MATHEMATICAL AND  
PHYSICAL CHARACTER**

**VOL. CXLV.**

**LONDON:**

**PRINTED FOR THE ROYAL SOCIETY AND SOLD BY  
HARRISON AND SONS, LTD., ST. MARTIN'S LANE  
PRINTERS IN ORDINARY TO HIS MAJESTY.**

**JULY, 1934.**

**LONDON:**  
**HARRISON AND SONS, LTD., PRINTERS IN ORDINARY TO HIS MAJESTY.**  
**ST. MARTIN'S LANE.**

# CONTENTS.

## SERIES A. VOL. CXLV.

No. A 854—June 2, 1934.

	PAGE
Faults in a Material which Yields to Shear Stress while Retaining its Volume Elasticity. By G. I. Taylor, F.R.S. ....	1
The Emission of Electrons under the Influence of Chemical Action. Part III.—The Action of $\text{Cl}_2$ , $\text{Br}_2$ , $\text{I}_2$ , $\text{NOCl}$ , $\text{HCl}$ , $\text{N}_2\text{O}$ , and $\text{COS}$ on $\text{NaK}_2$ . By A. K. Denisoff and O. W. Richardson, F.R.S. ....	18
Two-Dimensional Oscillations in Divergent Jets of Compressible Fluid. By S. G. Hooker. Communicated by R. V. Southwell, F.R.S. ....	52
An Analysis of the Dipole Moments of Some Aromatic Compounds. By G. M. Bennett and S. Glasstone. Communicated by C. K. Ingold, F.R.S. ....	71
The Nature and Properties of Aluminosilicate Framework Structures. By W. H. Taylor. Communicated by W. L. Bragg, F.R.S. ....	80
Stress Calculation for Tubular Frameworks having Continuous Longitudinals. By J. B. B. Owen. Communicated by R. V. Southwell, F.R.S. ....	104
An Optical Examination of Thin Films. I.—The Optical Constants of Mercury. By L. Tronstad and C. G. P. Feachem. Communicated by E. K. Rideal, F.R.S. ....	115
An Optical Examination of Thin Films. II.—The Behaviour of Thin Films of Fatty Acids on Mercury. By C. G. P. Feachem and L. Tronstad. Communicated by E. K. Rideal, F.R.S. ....	127
The Atomic Work Function of Tungsten for Potassium. By R. C. Evans. Communicated by Lord Rutherford, O.M., F.R.S. ....	135
On the Rotation of the Plane of Polarization of Long Radio Waves. By A. L. Green and G. Builder. Communicated by T. H. Laby, F.R.S. ....	145
On the Electron Theory of Metals. By S. Schubin and S. Wonsowsky. Communicated by R. H. Fowler, F.R.S. ....	159
Statistical Measurements of Turbulence in the Flow of Air through a Pipe. By H. C. H. Townend. Communicated by G. I. Taylor, F.R.S. (Plates 1-3) ....	180
Experimental Investigation and Analysis of the Velocity Variations in Turbulent Flow. By L. F. G. Simmons and O. Salter. Communicated by G. I. Taylor, F.R.S. (Plates 4-7) ....	212
An Analysis of the Fine Structure of the $\alpha$ -particle Groups from Thorium C and of the Long Range Groups from Thorium C'. By W. B. Lewis and B. V. Bowden. Communicated by Lord Rutherford, O.M., F.R.S. ....	235
An Automatic Magnetic Field Stabiliser of High Sensitivity. By C. E. Wynn-Williams. Communicated by Lord Rutherford, O.M., F.R.S. ....	250

	PAGE
<b>X-Ray Examination of Certain Copper-Zinc Alloys at Elevated Temperatures.</b> By E. A. Owen and L. Pickup. Communicated by Sir William Bragg, O.M., F.R.S.	258
<b>The Theory of the Change in Resistance in a Magnetic Field.</b> By H. Jones and O. Zener. Communicated by R. H. Fowler, F.R.S.	268
<b>The Structure of the Molecule of Nitrogen Dioxide from a Study of its Infra-Red Absorption Spectrum.</b> By G. B. B. M. Sutherland. Communicated by T. M. Lowry, F.R.S.	278
<b>The Rate of Transformation in Aqueous Solution of Methylammonium Cyanate into Methylurea.</b> By C. C. Miller. Communicated by Sir James Walker, F.R.S.	288
<b>The Reaction of Methane and Oxygen Sensitized by Nitrogen Peroxide. Part I.—Thermal Ignition.</b> By R. G. W. Norrish and J. Wallace. Communicated by T. M. Lowry, F.R.S.	307
<b>Band Spectrum of Aluminium Chloride (AlCl<sub>3</sub>).</b> By B. N. Bhaduri and A. Fowler, F.R.S. (Plate 8)	321
<b>Investigations in the Infra-red Region of the Spectrum. Part X.—The Asymmetrical Molecule Nitrosyl Chloride, NOCl.</b> By C. R. Bailey and A. B. D. Cassie. Communicated by F. G. Donnan, F.R.S.	336
<b>The Quantum Theory of the Neutron.</b> By G. Temple. Communicated by Sir Arthur Eddington, F.R.S.	344

No. A 855—July 2, 1934.

<b>The Mechanism of Plastic Deformation of Crystals. Part I.—Theoretical.</b> By G. I. Taylor, F.R.S.	362
<b>The Mechanism of Plastic Deformation of Crystals. Part II.—Comparison with Observations.</b> By G. I. Taylor, F.R.S.	388
<b>The Strength of Rock Salt.</b> By G. I. Taylor, F.R.S.	405
<b>The Vertical Distribution of Ozone in the Atmosphere.</b> By F. W. P. Götz, A. R. Meetham and G. M. B. Dobson, F.R.S.	416
<b>The Energies of Alpha, Beta, and Gamma Rays.</b> By H. A. Wilson, F.R.S.	447
<b>Note on a New Transition Produced by Electron Impact in Helium.</b> By R. Whiddington, F.R.S., and H. Priestley. (Plate 9)	462
<b>Small Angle Inelastic Electron Scattering in Helium, Neon, and Argon.</b> By R. Whiddington, F.R.S., and J. E. Taylor. (Plate 10)	465
<b>The Viscosity of Strong Electrolytes Measured by a Differential Method.</b> By W. M. Cox and J. H. Wolfenden. Communicated by Sir Harold Hartley, F.R.S.	475
<b>The Constitution of Water in Different States.</b> By I. Ramakrishna Rao. Communicated by O. W. Richardson, F.R.S.	489
<b>The Specific Heat of Nickel and of some Nickel-Copper Alloys.</b> By K. E. Grew. Communicated by R. Whiddington, F.R.S.	509
<b>A Theory of the Electrical Breakdown of Solid Dielectrics.</b> By C. Zener. Communicated by R. H. Fowler, F.R.S.	523

	PAGE
Some Peculiarities in the Physical Properties of Iron-Aluminium Alloys. By C. Sykes and H. Evans. Communicated by W. L. Bragg, F.R.S. ....	529
The Dielectric Constants of Liquid Mixtures of Phenol—Water, Phenol— <i>m</i> -Cresol, Phenol—Aniline, and Phenol— <i>p</i> -Toluidine. By O. R. Howell and W. Jackson. Communicated by T. M. Lowry, F.R.S. ....	539
The Catalytic Properties and Structure of Metal Films. Part II.—The Electrical Condition of Platinum Films. By G. I. Finch and (Miss) A. W. Ikin. Communicated by W. A. Bone, F.R.S. (Plate 11) ....	551
A Quantitative Study of Pleochroic Haloes.—I. By G. H. Henderson and S. Bateson. Communicated by Lord Rutherford, O.M., F.R.S. (Plates 12–14) ....	563
A Quantitative Study of Pleochroic Haloes.—II. By G. H. Henderson and L. G. Turnbull. Communicated by Lord Rutherford, O.M., F.R.S. (Plate 15) ....	582
A New Method of Determining the Age of Certain Minerals. By G. H. Henderson. Communicated by Lord Rutherford, O.M., F.R.S. ....	591
The Thermal Conductivity of Air by a Parallel Plate Method. By E. O. Hercus and D. M. Sutherland. Communicated by T. H. Laby, F.R.S. (Plate 16) ....	599
The Probability of Inner Shell Ionization of Atoms by Electron Impact. By E. H. S. Burhop. Communicated by T. H. Laby, F.R.S. ....	612
An Elementary Discussion of Ferromagnetism. By F. Bitter. Communicated by R. H. Fowler, F.R.S. ....	629
A Relativistic Basis of the Quantum Theory.—II. By H. T. Flint. Communicated by O. W. Richardson, F.R.S. ....	645
On the Polarization of Electrons by Scattering.—II. By E. G. Dymond. Communicated by C. G. Barkla, F.R.S. ....	657
On the Fracture of Fibred Iron-Silicon Sheets. By F. Bitter. Communicated by G. I. Taylor, F.R.S. (Plate 17) ....	668
The Beilby Layer. By G. I. Finch, A. G. Quarrell and J. S. Roebuck. Communicated by W. A. Bone, F.R.S. (Plate 18) ....	676
Investigations on the Spectrum of Selenium. III.—Extension of Se III. By K. R. Rao and S. G. K. Murti. Communicated by A. Fowler, F.R.S. ....	681
Investigations on the Spectrum of Selenium. IV.—Se I and Se VII. By K. R. Rao and S. G. K. Murti. Communicated by A. Fowler, F.R.S. ....	694
The Effect of Thermal Agitation on Atomic Arrangement in Alloys. By W. L. Bragg, F.R.S., and E. J. Williams. Forming the subject of the Bakerian Lecture by W. L. Bragg, F.R.S. ....	699
Index .....	733



# PROCEEDINGS OF THE ROYAL SOCIETY

## SECTION A—MATHEMATICAL AND PHYSICAL SCIENCES

### *Faults in a Material which Yields to Shear Stress while Retaining its Volume Elasticity.*

By G. I. TAYLOR, F.R.S., Yarrow Research Professor of the Royal Society.

(Received February 16, 1934.)

The theory of v. Mises that plastic flow begins when a certain quadratic function of the principal stresses reaches a certain value has been found to hold with materials like copper, aluminium, pure iron, nickel and certain types of mild steel.\* According to this theory the ratio

$$R = \frac{\text{shear yield stress in torsion}}{\text{tensile yield stress}}$$

should be  $1/\sqrt{3}$  or  $R = 0.577$ , and tests with tubes made of the material above mentioned do in fact give this result. The experiments of Cook† and others on steel tubes show that for some steels the observed value of  $R$  is very close to 0.50, which is the value it would have according to Mohr's maximum stress difference hypothesis.

Two reasons have been suggested for this very marked discrepancy between the results of experiments with different materials. In the first place with soft metals it is sometimes difficult to observe any definite yield point at which elastic failure occurs so that the most reliable results are then obtained by extrapolation from measurements in which a considerable amount of plastic flow had already occurred. v. Mises' law might therefore be a law of plastic flow rather than of elastic breakdown. The steels used by Cook had a very definite and easily observed yield point both in tension and torsion so that so

\* Taylor and Quinney, 'Phil. Trans.,' A, vol. 230, p. 323 (1932).

† 'Proc. Roy. Soc.,' A, vol. 137, p. 559 (1932).



far as these experiments are concerned, Mohr's maximum stress difference hypothesis may be a general law governing elastic breakdown in complex stress distributions.

The second possible explanation is that the steels for which the ratio  $R$  was found experimentally to be 0.50 were all steels which suffered a marked drop in stress at the yield point, while the materials for which  $R$  was found to be nearly 0.577 did not possess this characteristic. It may therefore be true in general that the law of elastic breakdown is that of Mohr for materials for which the stress drops at the yield point and that of v. Mises for materials which do not possess this property. One of the objects of this paper is to enquire whether there is any reason which would lead one to suspect that this may be so.

First consider the mechanism of breakdown in a material the stress of which does not drop at the yield point. The grains are initially self-stressed, the average stress being zero. When any stress is applied and gradually increased it might be expected that the greatest stress would occur in those grains which had initially an internal stress of the same type as the applied stress. These grains would be the first to yield and after yielding the stress in them would still increase with increasing average stress, but not so rapidly as that in the grains which were still in elastic strain. This process would rapidly bring about a condition in which the stress in each grain is the same as that of the whole mass so that the law of plastic yielding of each grain would be identical with that of the whole mass. It seems, therefore, that when v. Mises' law of plastic yielding is observed in metals which have no drop in stress at the yield point we are really observing in the mass a condition which applies separately to each grain and is a property of the metal itself as distinct from its granular structure.

Next let us consider the mechanism of elastic breakdown in materials the stresses in which suddenly decrease at the yield point.\* In this case, as in the previous one, the most stressed grains must be the first to yield, but when they yield the release of the stress which they were withstanding throws a heavier stress on the surrounding grains, some of which will yield in their turn. The material will yield as a whole when on the average each grain which yields throws such strain on the surrounding grains that at least one further grain yields without increase in mean stress. The effect of yielding of successive grains is therefore to make the stress distribution increasingly heterogeneous.

\* This includes all metals for which the stress in each grain drops as it yields, even though no drop in stress can be observed in the material as a whole.

This is the converse of the effect of yielding in a material which does not suffer a drop in stress at the yield point.

The condition of yielding of the material in bulk depends partly on the condition of yielding of the grains themselves and partly on the manner in which the yielding of one grain concentrates stress on its neighbours. From observations on the condition of yielding of material in bulk each grain of which has a drop in stress at the yield point, we can therefore make no direct deduction about the yielding of the individual grains. If we knew the shape of the grains and were able to carry out the necessary analysis we might be able to determine by means of the theory of elasticity the connection between the yielding of the grains and that of the material in bulk. Though this ideal is quite unattainable it seems possible to gain some insight into the problem by considering the effect on stress distribution in the surrounding metal of the yielding of grains of special forms for which the necessary calculations can be carried out.

In order to put the problem in a mathematically definable form it is necessary to make some simple assumption about the change which occurs in the stress inside a grain when it yields. The assumption which will be made is that at the instant a grain yields all shear stresses are released but the compressibility remains unaltered, so that any change in the volume of the grain which may result from its plastic deformation is accompanied by a corresponding change in the hydrostatic pressure of the material which has deformed. Thus from the point of view of stress distribution the grain becomes virtually a liquid possessing the same compressibility as the elastic material surrounding it.

The yield point of the material as a whole will be assumed to correspond to the value of the average stress which is just capable of raising the stress, at the point of maximum stress concentration, to such a value that local yielding begins there. We thus distinguish between the conditions of yielding of each element of the material and that of the whole mass. It will be noticed that the model which has been chosen possesses the property that if the yield condition of each element is independent of the hydrostatic pressure the yield condition of the whole mass is also independent of hydrostatic pressure. This condition would obviously not have been satisfied if we had used as our model an elastic material containing an empty hollow to represent the region where a grain had yielded.

We now proceed to consider two special cases :—

- (1) The material which yields fills a sphere.
- (2) It fills an elliptic cylinder.

In order to see the effect of granular structure on the yield condition we shall find the value in each case of the ratio

$$R = \frac{S}{P} = \frac{\text{shear yield stress in torsion}}{\text{tensile yield stress}},$$

making the alternative assumptions that the yield condition of each element of the mass is (a) that of v. Mises and (b) the maximum stress difference condition of Mohr. If the yield condition of the granular mass is identical with that of each elementary grain we should find  $R = 1/\sqrt{3} = 0.577$  for (a) and  $R = 0.500$  for (b).

*Yielding through a Spherical Volume.*

The equations for the elastic displacement due to a spherical hole in a material which is subject to a simple shearing strain have been given by Love.\* If the system be referred to axes parallel to the principal axes of stress, the displacements due to an extension  $\gamma$  per unit length parallel to  $ox$  and a contraction  $\gamma$  per unit length parallel to  $oy$  are:—

$$\left. \begin{aligned} u &= \gamma \left[ x + \frac{2x}{9\lambda + 14\mu} \left\{ 3(\lambda + \mu) \frac{a^5}{r^5} + 5\mu \frac{a^3}{r^3} \right\} \right. \\ &\quad \left. - 15 \frac{a^2 - r^2}{r^7} \left( \frac{\lambda + \mu}{9\lambda + 14\mu} \right) a^3 x (x^2 - y^2) \right] \\ v &= -\gamma \left[ y + \frac{2y}{9\lambda + 14\mu} \left\{ 3(\lambda + \mu) \frac{a^5}{r^5} + 5\mu \frac{a^3}{r^3} \right\} \right. \\ &\quad \left. + 15 \frac{a^2 - r^2}{r^7} \left( \frac{\lambda + \mu}{9\lambda + 14\mu} \right) a^3 y (x^2 - y^2) \right] \\ w &= -\gamma \left[ 15 \frac{a^2 - r^2}{r^7} \left( \frac{\lambda + \mu}{9\lambda + 14\mu} \right) a^3 z (x^2 - y^2) \right] \end{aligned} \right\}, \quad (1)$$

where  $a$  is the radius of the sphere  $r^2 = x^2 + y^2 + z^2$  and  $\lambda$  and  $\mu$  are the two elastic constants. The volume of the hollow remains unchanged by this strain so that if it were filled with liquid of the same compressibility as the elastic material the pressure inside would remain zero. At the surface of the sphere  $r = a$  the strain components are given by

$$\left. \begin{aligned} \frac{e_{xx}}{K\gamma} &= (\lambda + 2\mu) \left( 1 - \frac{2x^2}{a^2} \right) + 2(\lambda + \mu) \frac{x^2(x^2 - y^2)}{a^4} \\ \frac{e_{xy}}{K\gamma} &= -(\lambda + 2\mu) \left( 1 - \frac{2y^2}{a^2} \right) + 2(\lambda + \mu) \frac{y^2(x^2 - y^2)}{a^4} \end{aligned} \right\}, \quad (2)$$

\* "Theory of Elasticity," 4th ed., p. 252.

$$\left. \begin{aligned} \frac{e_{xx}}{K\gamma} &= 2(\lambda + \mu) x^2 \frac{(x^2 - y^2)}{a^4} \\ \frac{e_{yy}}{K\gamma} &= 2(\lambda + 2\mu) \frac{yz}{a^3} + 4(\lambda + \mu) yz \frac{(x^2 - y^2)}{a^4} \\ \frac{e_{zz}}{K\gamma} &= -2(\lambda + 2\mu) \frac{zx}{a^3} + 4(\lambda + \mu) zx \frac{(x^2 - y^2)}{a^4} \\ \frac{e_{xy}}{K\gamma} &= 4 \frac{xy(x^2 - y^2)}{a^4} (\lambda + \mu) \end{aligned} \right\}, \quad (2) \quad (\text{cont.})$$

where

$$K = \frac{15}{9\lambda + 14\mu}.$$

From the first three of these

$$\frac{\Delta}{K\gamma} = -2\mu \left( \frac{x^2 - y^2}{a^3} \right),$$

where

$$\Delta = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}.$$

We are now in a position to find the mean stress at which the yield point is first attained at some point in the field.

(a) Take first case (a) where the material yields when v. Mises' function

$$F = (\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2, \quad (3)$$

first attains a given value  $F$ . In this expression  $\sigma_1, \sigma_2, \sigma_3$  are the principal stresses. Expressed in terms of the strains and elastic constants

$$F = 4\mu^2 [(e_{xx} - e_{yy})^2 + (e_{yy} - e_{zz})^2 + (e_{zz} - e_{xx})^2] + 6\mu^2 [e_{xy}^2 + e_{yz}^2 + e_{zx}^2]. \quad (4)$$

Consider the values of  $F$  over the surface of the sphere. The form of the expressions for the strain components shows that  $F$  has a maximum or minimum value at the end of 3 diameters parallel to the principal stresses. At the point  $(a, 0, 0)$

$$e_{xx} = K\gamma\lambda, \quad e_{yy} = -K\gamma(\lambda + 2\mu), \quad e_{zz} = e_{yz} = e_{zx} = e_{xy} = 0.$$

Substituting these values in (4) the value of  $F$  is

$$4\mu^2 K^2 \gamma^2 (6\lambda^2 + 12\lambda\mu + 8\mu^2).$$

At the point  $(0, 0, a)$

$$e_{xx} = K\gamma(\lambda + 2\mu), \quad e_{yy} = -K\gamma(\lambda + 2\mu), \quad e_{zz} = 0,$$

so that

$$F = 4\mu^2 K^2 \gamma^2 (6\lambda^2 + 24\lambda\mu + 24\mu^2). \quad (5)$$

At the point  $(o, a, o)$   $F$  is the same as at  $(a, o, o)$ . The greatest value of  $F$ , therefore, occurs at the point  $(o, o, a)$  where the axis of  $z$  cuts the sphere. The position of this point is shown graphically in the diagrams, figs. 1 (a) and 1 (b).

To relate  $\gamma$  with  $S$  the shear stress in the material at great distances from the hole, notice that  $\Delta = 0$  so that  $S = 2\mu\gamma$ . Hence from (5) the value of  $S$  which raises the value of  $v$ . Mises' function to the critical value  $F$  at the point where it is greatest is given by

$$F = S^2 K^2 (6\lambda^2 + 24\lambda\mu + 24\mu^2). \quad (6)$$

When the material is subjected to a pure tensile stress  $P$  parallel to the axis of  $x$  the hole would increase in volume on applying the load but by adding a hydrostatic pressure equal to  $1/3P$  the sum of the three principal stresses is zero. The applied stress is then the sum of two pure shear stresses. The strain at great distances from the hole is also the sum of two pure shearing strains parallel to the planes  $xy$ ,  $xz$  respectively and the volume of the strained hole is equal to that of the hole before straining.

If the three principal strains at great distances from the hole are

$$e_{xx} = 2\gamma, \quad e_{yy} = -\gamma, \quad e_{zz} = -\gamma,$$

the stress components at  $(a, o, o)$  are

$$e_{xx} = 2\lambda K\gamma, \quad e_{yy} = e_{zz} = -K\gamma(\lambda + 2\mu), \quad e_{xy} = e_{xz} = e_{yz} = 0,$$

so that

$$F = 4\mu^2 K^2 \gamma^2 (18\lambda^2 + 24\lambda\mu + 8\mu^2).$$

At  $(o, a, o)$

$$e_{xx} = 2K\gamma(\lambda + 2\mu), \quad e_{yy} = -K\gamma\lambda, \quad e_{zz} = -K\gamma(\lambda + 2\mu),$$

so that

$$F = 4\mu^2 K^2 \gamma^2 (18\lambda^2 + 60\lambda\mu + 56\mu^2). \quad (7)$$

The greatest value of  $F$  occurs, therefore, on the equator, that is on the diametral plane perpendicular to the direction of greatest tension. The position of this equator is shown graphically in fig. 1 (c). From the elastic equation

$$X_x = 2P/3 = \lambda\Delta + 2\mu e_{xx} = 4\mu\gamma,$$

it will be seen that

$$P = 6\mu\gamma,$$

so that the tensile stress  $P$  which will just cause the material to yield at the point of greatest stress concentration is given by

$$F = \frac{4}{27} P^2 K^2 \mu^2 (18\lambda^2 + 60\lambda\mu + 56\mu^2). \quad (8)$$

The shear yield stress may be compared with the tensile yield stress by dividing (8) by (6) thus

$$\frac{S^2}{P^2} = \frac{18\lambda^2 + 60\lambda\mu + 56\mu^2}{9(6\lambda^2 + 24\lambda\mu + 24\mu^2)}. \quad (9)$$

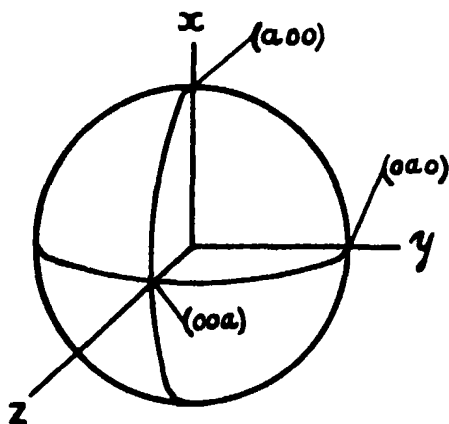


FIG. 1 (a).

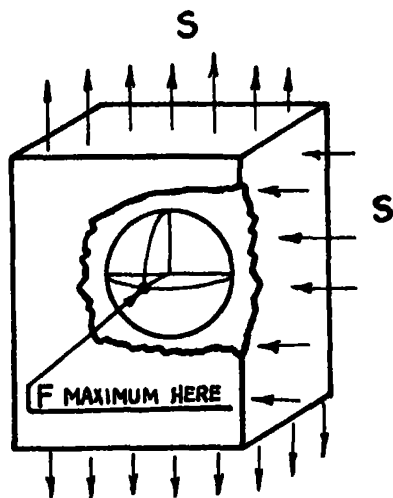


FIG. 1 (b).

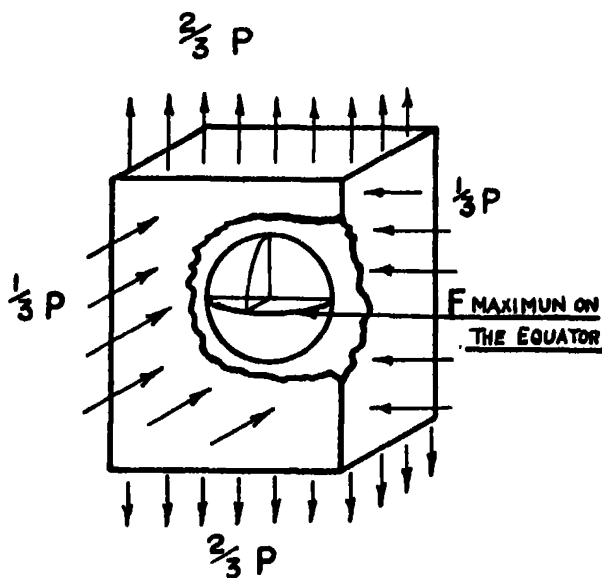


FIG. 1 (c).

Remembering that Poisson's ratio  $\sigma$  is equal to  $\frac{1}{2}\lambda/(\lambda + \mu)$  it will be found that (9) gives  $S/P = 0.525$  when  $\sigma = \frac{1}{4}$ ,  $S/P = 0.536$  when  $\sigma = 1/3$ , and

$S/P = 1/\sqrt{3}$  when  $\sigma = \frac{1}{2}$ . Comparing these with the value  $S/P = 1/\sqrt{3}$  applicable to the homogeneous material it will be seen that if the yielding of each element of the material is determined by v. Mises' law the yielding of the granular medium containing spherical holes will also be determined by v. Mises' law provided  $\sigma = \frac{1}{2}$ . For any other value of  $\sigma$  this is no longer true. For materials for which  $\sigma = \frac{1}{2}$ , for instance, the value 0.525 is nearer to 0.500 than to 0.577, so that the condition which determines yielding is nearer to that of Mohr than that of v. Mises.

(b) The maximum stress difference at the point of greatest stress concentration is found to be  $2SK (\lambda + 2\mu)$  for a pure shear stress  $S$ , and  $PK (\lambda + 2\mu)$  for a pure tensile stress  $P$ , so that if each element of the material is subject to Mohr's law that yielding begins when the maximum stress difference rises to a certain value, then  $S/P = \frac{1}{2}$ . Thus it appears that when each element obeys Mohr's law the granular mass also obeys Mohr's law.

### *Grains in the Form of Elliptic Cylinders.*

*Empty Elliptic Cavity.*—The distribution of stress round an elliptic cylindrical cavity in a material in a state of plane stress or plane strain was first found by Inglis.\* It was afterwards rediscovered by Pöschl† who gave an expression for the stress function due to a hole whose major axis is at any given angle  $\frac{1}{2}\pi + \alpha$  to the direction of a simple tensile stress  $p$ . Using elliptic co-ordinates derived from the transformation

$$\left. \begin{aligned} z = x + iy &= c \cosh (\xi + i\eta) = c \cosh \zeta \\ \text{so that} \quad x &= c \cosh \xi \cos \eta \\ y &= c \sinh \xi \sin \eta \end{aligned} \right\}. \quad (10)$$

Pöschl's expression for the stress function is

$$\psi = \frac{pc^2}{8} \{ \sinh 2\xi - \cos 2\alpha e^{-2(\xi-\xi_0)} - 2(\cos 2\xi_0 + \cos 2\alpha) \xi + [\cosh 2(\xi - \xi_0) - 1] e^{2\xi_0} \cos 2(\eta - \eta_0) \}, \quad (11)$$

where  $\xi = \xi_0$  is the equation to the elliptic hole. From this Pöschl deduces the stress  $\widehat{\eta\eta}$  round the surface of the hole. His expression‡ is

$$\widehat{\eta\eta} = p \frac{\sinh 2\xi_0 - \cos 2\alpha + e^{2\xi_0} \cos 2(\eta - \alpha)}{\cosh 2\xi_0 - \cos 2\eta}. \quad (12)$$

\* 'Proc. Inst. Naval Architects,' March 14, 1913.

† 'Math. Z.,' vol. 11, p. 89 (1921).

‡ There is a small misprint in Pöschl's paper but this is corrected in (12).

Consider now the stress distribution when the material is subjected to a simple shear so that at great distances from the hole  $X_{\infty} = p$ ,  $Y_{\infty} = -p$ . At the surface of the hole

$$\bar{\eta}\bar{\eta} = 2p \left[ \frac{-\cos 2\alpha + e^{2\xi_0} \cos 2(\eta - \alpha)}{\cosh 2\xi_0 - \cos 2\eta} \right]. \quad (13)$$

The stresses which must be applied to give the distributions represented by (12) and (13) are shown graphically in figs. 2 (a) and 2 (b).

For any given value of  $\alpha$  it is possible to find the maximum value of  $\bar{\eta}\bar{\eta}$  but in order that our model may represent a granular structure which is isotropic we must imagine that it contains elliptic grains placed at every possible angle to the principal axes of stress. Thus to find the maximum possible value of  $\bar{\eta}\bar{\eta}$  which can occur anywhere in the field we must find the maximum value of  $\bar{\eta}\bar{\eta}$  when both  $\eta$  and  $\alpha$  are allowed to vary. The corresponding values of  $\eta$  and  $\alpha$  are then given by the simultaneous equations

$$\frac{\partial}{\partial \eta} (\bar{\eta}\bar{\eta}) = 0 \quad \text{and} \quad \frac{\partial}{\partial \alpha} (\bar{\eta}\bar{\eta}) = 0. \quad (14)$$

Differentiating (13) it will be found that these reduce to

$$\left. \begin{aligned} \tan 2\alpha [\cosh 2\xi_0 \cos 2\eta - 1] + \sinh 2\xi_0 \sin 2\eta &= 0 \\ \tan 2\alpha [1 - e^{2\xi_0} \cos 2\eta] + e^{2\xi_0} \sin 2\eta &= 0 \end{aligned} \right\}. \quad (15)$$

The only possible solutions of these equations are  $\alpha = 0$  or  $\frac{1}{2}\pi$  and  $\eta = 0$  or  $\frac{1}{2}\pi$ , so that if elliptic holes of a given shape are distributed with their axes at random orientations, the maximum stresses occur at the ends of the major axes of those whose axes are parallel and perpendicular to the principal axes of stress. This might be interpreted to mean that if the material is full of cracks the cracks will extend in a direction perpendicular to the tensile stress, i.e., at  $45^\circ$  to the direction of the shearing stress. The orientations of the holes for maximum stress are shown in fig. 2 (c).

*Yielding over Elliptic Area.*—It has already been pointed out that the concentration of stress due to an empty hole cannot in general be similar to that resulting from the yielding of the material to shear stress while still retaining its elastic resistance to hydrostatic compression or expansion. In the particular case of a spherical hole in a material subject to shear stresses only, the volume of the hole is unaltered so that the pressure of the material within it is zero and it is immaterial whether we imagine the hole to be filled or empty. This does not obtain for the elliptic hole so that Pöschl's expression (13) cannot



be applied directly to find the stress concentrations due to local yielding. On the other hand we can calculate the change in the volume of the hole in Pöschl's case due to the applied stress, and can then apply a pressure uniform in all directions till the hole regains its original volume. The hole can then be regarded as being full of the material which has yielded and is still at zero pressure. Since further changes in hydrostatic pressure of the whole system

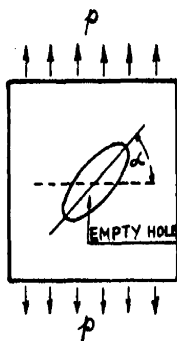


FIG. 2 (a).

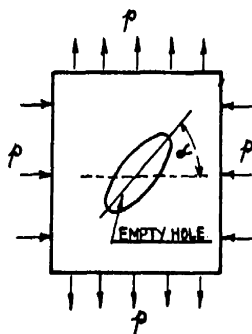


FIG. 2 (b).

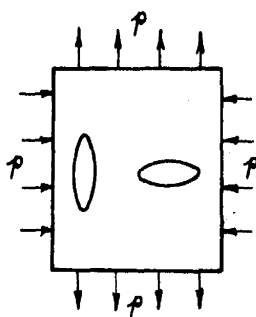


FIG. 2 (c).

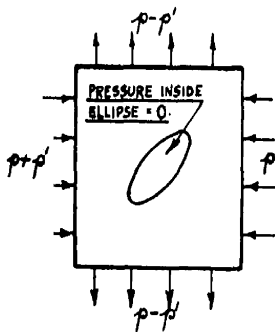


FIG. 2 (d).

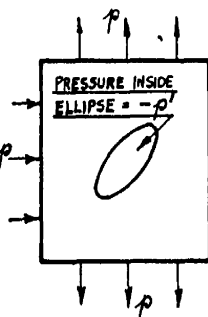


FIG. 2 (e).

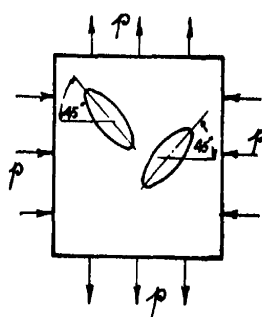


FIG. 2 (f).

do not then affect the stress differences, or shear stresses, we can calculate in this way as a problem in plane strain the distribution of shear stress round an elliptic grain which has yielded.

#### *Calculation of Change in Volume.*

We will first take Pöschl's case of a simple stress in one direction and calculate the change in volume of the elliptic hole. For this purpose it is necessary to

find the displacements. A general method for doing this when the stress function is known is given by Love.\* To apply this method we first find

$$- = 2(\lambda + \mu) \left( \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right).$$

The co-ordinates  $\xi, \eta$  are related to  $x, y$  by the relation

$$x + iy = c \cosh (\xi + i\eta) = c \cosh \zeta, \quad (16)$$

and the modulus of transformation is

$$h = \frac{2}{\sqrt{c^2 (\cosh 2\xi - \cos 2\eta)}},$$

so that

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = h^2 \left( \frac{\partial^2 \psi}{\partial \xi^2} + \frac{\partial^2 \psi}{\partial \eta^2} \right).$$

Applying this to Pöschl's expression (11)

$$\Delta = \frac{p}{2(\lambda + \mu) (\cosh 2\xi - \cos 2\eta)} [\sinh 2\xi (1 + e^{2\epsilon_0} \cos 2\alpha) + e^{2\epsilon_0} \sin 2\alpha \sin 2\eta + e^{2\epsilon_0} \cos 2\alpha \{-\cosh 2\xi + \cos 2(\eta - \alpha)\}]. \quad (17)$$

Next find  $2i\bar{w}$  (where  $\bar{w}$  is the rotation) from the equations for equilibrium and dilatation, namely,

$$(\lambda + 2\mu) \frac{\partial \Delta}{\partial x} - 2\mu \frac{\partial \bar{w}}{\partial y} = 0, \quad (\lambda + 2\mu) \frac{\partial \Delta}{\partial y} + 2\mu \frac{\partial \bar{w}}{\partial x} = 0,$$

and express  $(\lambda + 2\mu) \Delta + 2i\mu\bar{w}$  as a function of  $\zeta$ . It will be found

$$(\lambda + 2\mu) \Delta + 2i\mu\bar{w} = \frac{p(\lambda + 2\mu)}{2(\lambda + \mu)} [(1 + e^{2\epsilon_0} \cos 2\alpha + ie^{2\epsilon_0} \sin 2\alpha) \coth \zeta - e^{2\epsilon_0} \cos 2\alpha]. \quad (18)$$

The function  $\Xi + iH$  is then defined by the equation

$$\begin{aligned} \Xi + iH &= \int \{(\lambda + 2\mu) \Delta + 2i\mu\bar{w}\} \frac{dz}{d\zeta} d\zeta \\ &= \frac{p(\lambda + 2\mu)}{2(\lambda + \mu)} c \{ (1 + e^{2\epsilon_0} \cos 2\alpha + ie^{2\epsilon_0} \sin 2\alpha) \sinh \zeta \\ &\quad - e^{2\epsilon_0} \cos 2\alpha \cosh \zeta \}. \quad (19) \end{aligned}$$

\* "Mathematical Theory of Elasticity," 4th ed., p. 204.

The displacements  $u, v$  parallel to the axes  $x$  and  $y$  are then given by

$$2\mu u = -\frac{\partial\psi}{\partial x} + \Xi, \quad 2\mu v = -\frac{\partial\psi}{\partial y} + \mathbf{H}.$$

The displacements  $u_\xi, u_\eta$  parallel to  $\xi$  and  $\eta$  are

$$2\mu u_\xi = -\hbar \frac{\partial\psi}{\partial \xi} + \mathbf{R} \left[ \frac{\Xi + i\mathbf{H}}{\hbar c \sinh \zeta} \right]$$

$$2\mu u_\eta = -\hbar \frac{\partial\psi}{\partial \eta} + \mathbf{I} \left[ \frac{\Xi + i\mathbf{H}}{\hbar c \sinh \zeta} \right],$$

where  $\mathbf{R}$  and  $i\mathbf{I}$  represent the real and imaginary parts.

Taking the real part of  $\frac{\Xi + i\mathbf{H}}{\sinh \zeta}$  and differentiating Pöschl's expression for  $\psi$ , it is found that

$$u_\xi = -\frac{pc^2\hbar}{8\mu} [\cosh 2\xi + \cos 2\alpha e^{-2(\xi - \xi_0)} - (\cosh 2\xi_0 + \cos 2\alpha) \\ + \sinh 2(\xi - \xi_0)e^{2\xi_0} \cos 2(\eta - \alpha)] \\ + \frac{p(\lambda + 2\mu)}{4\mu\hbar(\lambda + \mu)} \left[ 1 + e^{2\xi_0} \cos 2\alpha - \frac{c^2\hbar^2}{2} e^{2\xi_0} \cos 2\alpha \sinh 2\xi \right] \quad (20)$$

$$u_\eta = -\frac{pc^2\hbar}{8\mu} \cos 2[(\xi - \xi_0) - 1] e^{2\xi_0} \sin 2(\eta - \eta_0) \\ + \frac{p(\lambda + 2\mu)}{4\mu\hbar(\lambda + \mu)} \left[ e^{2\xi_0} \sin 2\alpha + \frac{c^2\hbar^2}{2} e^{2\xi_0} \cos 2\alpha \sin 2\eta \right]$$

The change in volume of the hole is  $\int_0^{2\pi} \frac{u_{\xi_0}}{\hbar} d\eta$ ,

where  $u_{\xi_0}$  is the value of  $u_\xi$  at  $\xi = \xi_0$ . From (20) it will be found that

$$\frac{1}{\hbar} u_{\xi_0} = \frac{p(\lambda + 2\mu)c^2}{8\mu(\lambda + \mu)} [\cos 2\alpha + \cosh 2\xi_0 - (1 + e^{2\xi_0} \cos 2\alpha) \cos 2\eta], \quad (21)$$

so that the change in volume due to a stress  $p$  acting at angle  $\pi/2 + \alpha$  to the major axis of the ellipse is

$$\frac{\sigma^2 p}{8\mu} \left( \frac{\lambda + 2\mu}{\lambda + \mu} \right) 2\pi (\cos 2\alpha + \cosh 2\xi_0). \quad (22)$$

Replacing  $\alpha$  by  $\alpha - \pi/2$  and  $p$  by  $-p$  in (22) the increase in volume of the hole due to a stress  $-p$  acting at angle  $\alpha$  is

$$\frac{\sigma^2 p\pi}{4\mu} (\cos 2\alpha - \cosh 2\xi_0), \quad (23)$$

and adding (22) and (23) it will be seen that the change in volume due to a pure shear stress formed by combining the two principal stresses  $p$  at angle  $\pi/2 + \alpha$  and  $-p$  at angle  $\alpha$  is

$$\frac{\pi \omega^2 p}{2\mu} \left( \frac{\lambda + 2\mu}{\lambda + \mu} \right) \cos 2\alpha. \quad (24)$$

The change in volume due to uniform pressure  $p'$  acting in all directions in the plane  $xy$  can be found by subtracting (22) from (23). It is

$$-\frac{\pi \omega^2 p'}{2\mu} \left( \frac{\lambda + 2\mu}{\lambda + \mu} \right) \cosh 2\xi_0. \quad (25)$$

In order that the change in volume due to the combination of the shear stress ( $p, -p$ ) with the uniform pressure ( $p', p'$ ) may be zero it is therefore necessary that

$$p \cos 2\alpha - p' \cosh 2\xi_0 = 0,$$

so that

$$\frac{p'}{p} = \frac{\cos 2\alpha}{\cosh 2\xi_0}. \quad (26)$$

### *Calculation of Stress Distribution due to Yielding inside the Ellipse.*

We are now in a position to calculate the stress distribution round an elliptic portion of material which has yielded in shear but not in uniform compression. The greatest stresses are likely to occur on the ellipse itself. Pöschl's expression for the stress  $\widehat{\eta\eta}$  at the surface due to a single stress  $p$  at angle  $\pi/2 + \alpha$  is given in (12). For the shearing stress ( $p$  at angle  $\pi/2 + \alpha, -p$  at angle  $\alpha$ )

$$\widehat{\eta\eta} = 2p \left[ \frac{-\cos 2\alpha + e^{2\xi_0} \cos 2(\eta - \alpha)}{\cosh 2\xi_0 - \cos 2\eta} \right]. \quad (27)$$

For a uniform tension  $p'$  applied equally in all directions in the plane

$$\widehat{\eta\eta} = 2p' \frac{\sinh 2\xi_0}{\cosh 2\xi_0 - \cos 2\eta}. \quad (28)$$

Hence for the combination which does not change the volume of the hole it is necessary to add (27) and (28) at the same time replacing  $p'$  by

$$p \frac{\cos 2\alpha}{\cosh 2\xi_0}.$$

The resulting expression

$$\widehat{\eta\eta} = 2p \left[ \frac{-\cos 2\alpha (1 + \tanh 2\xi_0) + e^{2\xi_0} \cos 2(\eta - \alpha)}{\cosh 2\xi_0 - \cos 2\eta} \right]. \quad (29)$$

The external stresses which give this value for the stress at the surface of the ellipse are shown graphically in fig. 2 (*d*).

When the externally applied stress is a pure shear stress ( $p, -p$ ) the pressure in the cavity is  $-p'$  but the shear stresses are the same as in the case represented in fig. 2 (*d*). This condition is shown in fig. 2 (*e*).

To find the maximum possible value of  $\bar{\eta}\eta$  when  $\alpha$  is allowed to assume any value we can proceed as for the empty elliptic hole, thus

$$\frac{\partial}{\partial \eta} \bar{\eta}\eta = 0, \quad (30)$$

and

$$\frac{\partial}{\partial \alpha} \bar{\eta}\eta = 0. \quad (31)$$

Differentiating (29) the equations for  $\eta$  and  $\alpha$  reduce to

$$\sin 2\alpha (\cosh 2\xi_0 \cos 2\eta - 1) = \cos 2\alpha \sin 2\eta \frac{\sinh^2 2\xi_0}{\cosh 2\xi_0}, \quad (32)$$

$$\sin 2\alpha (1 - \cosh 2\xi_0 \cos 2\eta) + \cosh 2\xi_0 \cos 2\alpha \sin 2\eta = 0. \quad (33)$$

The solution of these equations is

$$\left. \begin{aligned} \cos 2\alpha &= 0 \\ \cos 2\eta &= \frac{1}{\cosh 2\xi_0} \end{aligned} \right\}. \quad (34)$$

It appears, therefore, that for any given shape of ellipse, i.e., for any given value of  $\xi_0$ , the maximum value of  $\bar{\eta}\eta$  is attained when  $\alpha = \frac{1}{4}\pi$ , i.e., when the axes of the ellipse are at  $45^\circ$  to the directions of the principal stresses. Such faults are shown graphically in fig. 2 (*f*). The maximum value of  $\bar{\eta}\eta$  is then

$$\frac{4p}{1 - e^{-2\xi_0}},$$

or if the applied stress is expressed in terms of a shearing stress  $S$  the maximum value of  $\bar{\eta}\eta$  is

$$\frac{4S}{1 - e^{-2\xi_0}}. \quad (35)$$

In an elongated ellipse  $\xi_0$  is small and the maximum value of  $\bar{\eta}\eta$  is approximately  $2S/\xi_0$ . In this case it is of interest to see how the maximum value of  $\bar{\eta}\eta$  over the surface of an ellipse of given orientation  $\alpha$  varies with  $\alpha$ . Using only the equation  $\partial(\bar{\eta}\eta)/\partial\eta = 0$  it will be seen from (32) that when  $\xi_0$  is small  $\cos 2\eta = 1/\cosh 2\xi_0$  as before, and inserting this in the expression for  $\bar{\eta}\eta$  it

is found that the maximum value of  $\eta\eta$  is approximately  $2S \sin 2\alpha/\xi_0$ . Thus when the elongated ellipse lies with its axes parallel to the principal axes of stress so that  $\alpha = 0$  then  $\eta\eta = 0$  and there is no concentration of stress due to the yielding. Faults of this kind are shown in fig. 2 (c).

*Propagation of Faults and Luder's Lines.*

The result just obtained is a remarkable one. It brings out very clearly the profound difference which exists between the concentration of stress produced by an elongated empty hole and that produced by failure of shear stress in an elongated volume without change in resistance to compression or expansion. In the former case failure is propagated as a crack running perpendicular to the principal stress and in the latter as a fault running in a direction inclined at  $45^\circ$  to the directions of the principal stresses.

It seems that the hypothetical substance which forms the subject of the preceding analysis has a property which is possessed by all materials in which Luder's lines can be produced. Faults can propagate themselves at  $45^\circ$  to the principal axes of stress when the load is less than that necessary for yielding throughout the mass. These faults can start from any hole or groove where there is an initial concentration of stress and the release of shear stress without release of compressive stress ensures that the fault will propagate itself at  $45^\circ$  to the direction of the principal stresses.

*Yield Criterion as Condition under which Faults can be Propagated.*

It has been seen that the presence in a material of spherical volumes within which the shear stresses have disappeared causes the whole mass to yield according to a law which is nearer to Mohr's than v. Mises' hypothesis, even though each element of the material obeys v. Mises' law.

It will now be shown that the effect of replacing these spherical faults by faults occupying elongated elliptic cylinders is to accentuate this effect so that in the limit narrow faults will be propagated only when the maximum stress difference rises to a certain value irrespective of whether each element of the material itself obeys Mohr's or v. Mises' law.

Consider a fault in the form of an elliptic cylinder  $\xi = \xi_0$  with its axes at  $45^\circ$  to the principal stresses; a shear stress  $S$  produces a maximum value of  $\eta\eta$  equal to  $4S/(1 - e^{-2\epsilon})$ , the other principal stresses in plane strain are

$$\xi\xi = 0 \quad \text{and} \quad \hat{z}z = \frac{4\sigma S}{1 - e^{-2\epsilon}},$$

where  $\sigma$  is Poisson's ratio. Hence v. Mises' function is

$$F = \frac{32S^2}{(1 - e^{-2\xi_0})^2} (1 - \sigma + \sigma^2).$$

In the limit  $\xi_0 \rightarrow 0$  this becomes

$$F = 8S^2 \xi_0^{-2} (1 - \sigma + \sigma^2). \quad (36)$$

We can obtain the distribution of stress round the ellipse when a simple stress  $P$  is applied by starting with a shearing stress system  $(\frac{1}{2}P, -\frac{1}{2}P, 0)$ . With this system of applied stress the values of the three principal stresses at the point of maximum stress concentration are

$$\widehat{\eta\eta} = \frac{2P}{1 - e^{-2\xi_0}}, \quad \widehat{z z} = \frac{2\sigma P}{1 - e^{-2\xi_0}}, \quad \widehat{\xi\xi} = 0.$$

Next apply a uniform stress  $\frac{1}{2}P$  in all directions, and finally apply an additional stress  $-\frac{1}{2}P$  perpendicular to the plane of the shear stress. The stresses at the maximum concentration are then

$$\widehat{\xi\xi} = \frac{1}{2}P, \quad \widehat{\eta\eta} = \frac{2P}{1 - e^{-2\xi_0}} + \frac{1}{2}P, \quad \widehat{z z} = \frac{2\sigma P}{1 - e^{-2\xi_0}},$$

and at some distance from the crack they are

$$X_x = P, \quad Y_y = 0, \quad Z_z = 0.$$

v. Mises' function is then

$$F = 4P^2 \left[ \left( \frac{1 - \sigma}{1 - e^{-2\xi_0}} + \frac{1}{2} \right)^2 + \frac{1}{(1 - e^{-2\xi_0})^2} + \left( \frac{\sigma}{1 - e^{-2\xi_0}} - \frac{1}{2} \right)^2 \right].$$

In the limit when  $\xi_0 \rightarrow 0$  this is

$$F = \frac{8P^2}{4\xi_0^2} (1 - \sigma + \sigma^2). \quad (37)$$

Comparing (36) and (37) it will be seen that when  $\xi_0 \rightarrow 0$ , i.e., when the ellipses are very elongated

$$S^2/P^2 = \frac{1}{4} \quad \text{or} \quad S/P = \frac{1}{2}.$$

Thus when each element of the material fails according to v. Mises' law the material as a whole fails according to the maximum difference law of Mohr. It is obvious that if each element of the material fails according to Mohr's law the material as a whole does so also. It seems that when failure is by propagation of narrow faults the law of failure of the granular mass will be that of Mohr whatever the law of failure of individual grains may be.

Perhaps this result might have been anticipated without analysis, for the kind of fault which can give rise to large concentrations of stress is essentially two dimensional in character. The greatest stresses occur near its edge and the large accompanying strains are due to displacements in the plane perpendicular to that edge. The stress component parallel to the edge is only increased in proportion to the stresses in the plane perpendicular to the edge by an amount proportional to Poisson's ratio. There is no concentration of a stress externally applied to the direction parallel to the edge. The resulting concentration in the value of *v. Mises'* function depends therefore almost entirely on the components of stress in the plane perpendicular to the edge of the fault, thus the intermediate principal stress in the granular mass as a whole does not affect the maximum value of *v. Mises'* function at the point of greatest stress concentration in a fault.

#### *Summary and Conclusions.*

Plastic substances may be divided into two classes according to whether the stress increases or decreases after yielding. In the former the stress in the grains which have yielded increases more slowly with increasing strain than in those which are still elastic. The effect of successive yielding of grains is therefore to reduce the internal stresses to a state of uniformity. The yield condition of the whole granular structure is therefore identical with that of each grain. In materials which lose their power to withstand shear stresses when they yield, the yielding of successive grains might be expected to produce an increasingly heterogeneous distribution of internal stresses.

To trace the effect of this heterogeneity on the yield properties of the whole granular mass a hypothetical material is imagined which has the property that when any portion yields all shear stresses vanish but the compressibility remains unaltered. The distributions of stress in the neighbourhood of grains which have yielded when their boundaries are spheres and elliptic cylinders are discussed. It is shown that when each element of the material fails according to the law of *v. Mises* the effect of spherical faults is to make the whole mass fail according to a law which is intermediate between that of *v. Mises* and *Mohr*. Narrow elliptic faults on the other hand cause the whole mass to fail according to *Mohr's* law of maximum stress difference even though each element fails according to *v. Mises'* law. When the law of failure of each element is that of *Mohr* the law of failure of the whole mass is also that of *Mohr*.

The hypothetical material has the property that the greatest stress concentrations occur when elongated faults lie at  $45^\circ$  to the directions of the



principal stresses, so that faults once started would propagate themselves along lines at  $45^\circ$  to the principal stresses. This property seems to have its counterpart in real materials which can exhibit Luder's lines. The fact that the greatest concentration of stress occurs when a fault lies at  $45^\circ$  to the direction of the principal stresses is in striking contrast to the case of a crack or fault in which compressive or tensile stress is released as well as shear stress. In that case the greatest concentration of stresses occur when the fault is perpendicular to the direction of greatest tension so that faults of that type would be propagated in the direction of one of the principal stresses.

---

*The Emission of Electrons under the Influence of Chemical Action.*  
*Part III.—The Action of  $\text{Cl}_2$ ,  $\text{Br}_2$ ,  $\text{I}_2$ ,  $\text{NOCl}$ ,  $\text{HCl}$ ,  $\text{N}_2\text{O}$ , and  $\text{COS}$  on  $\text{NaK}_2$ .*

By A. K. DENISOFF, University of London, King's College, and  
O. W. RICHARDSON, F.R.S., Yarrow Research Professor of the Royal Society.

(Received February 1, 1934.)

§ 1. *Introduction.*

The two preceding papers,\* which will be referred to in the sequel as Part I and Part II, dealt chiefly with a detailed investigation of the electron emission when the liquid alloy of sodium and potassium of the composition  $\text{NaK}_2$  is acted on by phosgene ( $\text{COCl}_2$ ). The latter substance was chosen for this purpose mainly because it is comparatively easy to manipulate on account of its extremely low vapour pressure at the temperature of liquid air and almost atmospheric pressure at room temperatures. A number of the conclusions in Part II were based on the results of unpublished experiments with other gaseous reagents. Some of these are described in the present paper.

After the completion of most of the experiments described in Part I and Part II the further investigation of the phenomenon of chemical electron emission was directed towards clearing up (1) the effect of the mechanism of the chemical reaction; (2) the connection between the amount of energy available in the reaction and the energy of the electrons emitted; and (3) the

\* Denisoff and Richardson, 'Proc. Roy. Soc.,' A, vol. 132, p. 22 (1931), (Part I); *ibid.*, vol. 144, p. 46 (1934), (Part II).

changes caused by varying the nature of the molecules taking part in the reaction.

With this object three different groups of gases were tried :—

- (I)  $\text{Cl}_2$ ,  $\text{COCl}_2$ ,  $\text{NOCl}$ , and  $\text{HCl}$ .
- (II)  $\text{Cl}_2$ ,  $\text{Br}_2$ , and  $\text{I}_2$ .
- (III)  $\text{H}_2\text{O}$ ,  $\text{N}_2\text{O}$ , and  $\text{COS}$ .

It was known from older experiments that electron emission occurred when a number of these gases at low pressures reacted with the liquid alloys of sodium and potassium. Thus Haber and Just\* investigated  $\text{Br}_2$  and  $\text{I}_2$  as well as  $\text{COCl}_2$ . They also examined the action of  $\text{COCl}_2$  on caesium and on very dilute amalgams of lithium, potassium, and caesium. Their general impression, as justified by the results of their investigations, was to the effect that the magnitude of the reaction effect follows in parallel with the amount of the chemical energy consumed in the transformation. The carriers of the emission from the amalgams, although all of them were negatively charged, were found to be heavy ions and not electrons.

The distribution of the energy of the electrons emitted from sodium potassium alloys when acted on at low pressures by  $\text{Cl}_2$ ,  $\text{H}_2\text{O}$ , and  $\text{HCl}$ , as well as by  $\text{COCl}_2$ , was previously investigated by one of us.† For  $\text{Cl}_2$  and  $\text{H}_2\text{O}$  the characteristic curves were found to be of the same general nature as for  $\text{COCl}_2$  and there was no striking difference in the magnitudes of the observed emissions. The average energy of the emitted electrons was found to be somewhat higher for  $\text{Cl}_2$  than for  $\text{COCl}_2$ . With  $\text{HCl}$  the currents were much smaller than those given by  $\text{Cl}_2$  and  $\text{COCl}_2$ , too small, in fact, for satisfactory measurements to be made; the characteristic curves appeared to be similar to those for the other gases, the currents dropping from the saturation value to zero in a range of potential difference of the order of 1 volt. These old observations on water vapour are quite different from what has been found in the present research. We shall return to this point later.

## § 2. *The Methods of Representing the Observational Data.*

Our chief objective in the experiments has been the determination of the energy distribution, or energy spectrum, of the emitted electrons. The previous work with  $\text{COCl}_2$  has shown that this is independent of a great variety of experimental conditions and other circumstances and is evidently something

\* 'Ann. Physik,' vol. 36, p. 308 (1911).

† Richardson, 'Phil. Trans.,' A, vol. 222, p. 1 (1921).

which is characteristic of the reaction under consideration. Our results with other gases enable us to state that this result is true generally.

The first step in this direction is to ascertain the "characteristic curve" which represents the relation between the observed current  $i$  and the applied voltage  $V_1$  read on the voltmeter employed for the purpose. This is not, in general, the true potential difference  $V$  between the drop and the surrounding electrode, but is connected with  $V$  by the relation  $V = V_1 + K$ , where  $K$  is the contact potential difference between the alloy and the outer electrode. At sufficiently low pressures, as pointed out in Part II, the point on the  $V_1$  scale at which  $V = 0$  is the point at which the electronic currents just reach the maximum or saturation value. This conclusion has been confirmed by experiments on mixtures of gases which will be described in detail in a subsequent communication. The actual energy spectrum curves which express the proportion of electrons having kinetic energies within a given range of values, such as between  $eV$  and  $e(V + dV)$ , are then obtained by differentiating the characteristic curves (graphically) with respect to  $V$ .

While these two groups of curves suffice to describe important features of the phenomena they are inadequate to describe with sufficient accuracy the observations in the region which corresponds to the emission of electrons with higher energies. So long as the currents in the retarding electrostatic fields employed are greater than about  $10^{-13}$  amp. the absolute error  $\Delta i$  of the measurement of the current  $i$  was chiefly determined by the inconsistency  $\Delta T$  of the time  $T$  of the alkali metal drops. In our experiments with the retarding field method we measure  $i$  at different  $-V$ ; thus, if the change in  $T$  from  $T$  to  $T + \Delta T$  corresponds to a change in  $i$  from  $i$  to  $i + \Delta i$  then the absolute error of the measurement of the current  $i$  is

$$\Delta i = (\Delta i_0) \frac{i}{i_0}, \quad (1)$$

where  $i_0$  is the saturation current (at  $V = 0$ ). So that the relative (percentage) error  $\Delta i/i$  of the observations is largely independent of the magnitude of the current, but the absolute error  $\Delta i$  of the measurements is greatest for the largest currents. The smaller currents which have been measured with much the same relative accuracy as the others can only be expressed very roughly if at all on a diagram in which the scale of current is small enough to express the value of the saturation current. A very convenient way of overcoming this difficulty is to take a logarithmic scale for  $i/i_0$ , since on such a diagram the absolute error  $\Delta i$  will be independent of  $V$ .

The  $\log i/i_0$  vs.  $V$  (or  $V_1$ ) plots are, at the same time, very instructive in detecting a small group of electrons which has an average energy higher than that of the main group. As we shall see later, such a small group of high average energy electrons is present each time when the chemical reaction has a pronounced two-stage mechanism. This important property of the  $\log i/i_0$  vs.  $V$  plots is due to the fact that the electron energy distribution function is generally of the form  $N(V) dV = A \cdot e^{-k(V-V_m)^c} dV$ , where  $A$ ,  $k$ ,  $c$ , and  $V_m$  are constants; the constant  $c$  has a value very close to 2 for the more energetic and very close to 1 for the less energetic reactions (see § 1, Part II).

We shall first consider the case in which  $c \approx 1$ ; the electron current  $i$  is then proportional to  $e^{-kV}$ , i.e., the  $\log i/i_0$  vs.  $V$  plot will be a straight line, the slope of which is equal to  $k$ . The constant  $k$  may be regarded as a measure for the average electron energy—the larger the energy the smaller the constant  $k$ . Now, suppose that the electron spectrum consists of two electron groups with the total number of electrons  $i_1$  and  $i_2$ , so that the total current is  $i_0 = i_1 + i_2$ , and suppose that  $i_1 \ll i_2$ , and  $k_1$ , which refers to  $i_1$ , is  $< k_2$ . We can write

$$\frac{i}{i_0} = \frac{i_1}{i_0} e^{-k_1 V} + \frac{i_2}{i_0} e^{-k_2 V}. \quad (2)$$

$i_1$  will contribute very little to  $i$  in the neighbourhood of  $V = 0$ , and in this region the relation between  $i$  and  $V$  will be practically

$$\frac{i}{i_0} = \frac{i_2}{i_0} e^{-k_2 V}, \quad (3)$$

or

$$\log \frac{i}{i_0} = \log \frac{i_2}{i_0} - k_2 V, \quad (4)$$

so that the  $\log i/i_0$  vs.  $V$  plot will start from the origin ( $V = 0$ ) as a straight line whose slope is  $k_2$ . The current  $i_0$  here is  $= i_2$  since  $i_1/i_0$  is regarded as negligible. However, at very high potentials  $e^{-k_1 V}$  will always become very small compared with  $e^{-k_2 V}$ , consequently if  $V$  is high enough this will always ultimately outweigh the other factor, and for high enough values of  $V$  we shall have

$$\frac{i}{i_0} = \frac{i_1}{i_0} e^{-k_1 V}, \quad (5)$$

or

$$\log \frac{i}{i_0} = \log \frac{i_1}{i_0} - k_1 V, \quad (6)$$

this will give another straight line plot whose slope is  $k_1$  and whose ordinate at  $V = 0$  is  $\log i_1/i_0$ .

If the current  $i_1$ , resulting from the higher average energy group of electrons, is comparable in magnitude with  $i_2$  then for large values of  $V$  we shall have the same relation (6). From this we can find the values of  $\frac{i_1}{i_0} e^{-k_1 V}$  for lower values of  $V$  for which  $\frac{i_2}{i_0} e^{-k_2 V}$  is not small. If we subtract these values of  $\frac{i_1}{i_0} e^{-k_1 V}$  from the corresponding values of  $i/i_0$ , calling the resulting difference  $i'/i_0$ , i.e.,

$$\frac{i'}{i_0} = \frac{i}{i_0} - \frac{i_1}{i_0} e^{-k_1 V}, \quad (7)$$

we shall have from (2)

$$\frac{i'}{i_0} = \frac{i_2}{i_0} e^{-k_2 V}, \quad (8)$$

or

$$\log \frac{i'}{i_0} = \log \frac{i_2}{i_0} - k_2 V. \quad (9)$$

This should give another linear plot whose ordinate at  $V = 0$  is  $\log i_2/i_0$  and whose slope is  $k_2$ .

This method is, however, inadequate to detect a small group of lower average-energy electrons, i.e., when  $i_1 \gg i_2$ . In this case  $\frac{i_2}{i_0} e^{-k_2 V}$  will be small compared with  $\frac{i_1}{i_0} e^{-k_1 V}$  for all values of  $V$  from 0 to  $\infty$ . The experiments will not detect a small low energy emission because it will never contribute anything appreciable to  $i/i_0$ . Its maximum relative contribution is near  $V = 0$  where the absolute error of the measurements is greatest, and this will increase the difficulty of detecting this type of emission.

For the more energetic reactions, for which  $c \approx 2$ , the  $\log i/i_0$  vs.  $V$  plots will still be able to detect a small admixture with higher average energy than that of the main group, but no simple quantitative treatment, similar to the above, is possible for these electron spectra. Instead of a straight line the  $\log i/i_0$  vs.  $V$  plot will give a curve, convex towards the origin ( $V = 0$ ), the slope of which continuously increases as  $\log i/i_0$  increases. A small high energy group of electrons will make itself evident by the fact that at high enough values of  $V$  the systematic increase in the curvature will become less rapid and may even become negative. Thus, the  $\log i/i_0$  vs.  $V$  plot will be convex at low values of  $V$  and concave at high values of  $V$ .

It will be clear from what has been said that the observational data given in the paper include two sets of experimental records. These are the experimental points on the characteristic plots  $i/i_0$  vs.  $V_1$  and on the  $\log i/i_0$  vs.  $V_1$

plots respectively. The experimental points on the characteristic plots only reproduce the data to the accuracy of the observations in the range of  $i$  from  $i_0$  to about  $i_0/20$ . Lower values of  $i$  than this should be obtained from the  $\log i/i_0$  vs.  $V_1$  plots. The energy distribution curves have been obtained by graphical differentiation of the smooth characteristic curves and consequently are less accurate, as a record of observational data, than the experimental points on the  $i/i_0$  vs.  $V_1$  and  $\log i/i_0$  vs.  $V_1$  plots.

The subject matter of this section will be considered further in the next communication (Part IV) where it will also be tested by the results of experiments.

### § 3. *Experimental Arrangements.*

The present investigation of the electron emission under the action of various active gases on the liquid alloy of sodium and potassium was carried out with the apparatus of which a detailed description was given in Part I. As for  $\text{COCl}_2$ , the tube containing the gas under investigation, in the liquid state, was placed in the corresponding freezing mixture, thus making it possible to reduce the pressure of the saturated vapour of the gas down to a convenient magnitude, usually some millimetres. Always, in testing a new gas the vacuum taps were cleaned and the grease was changed in that part of the apparatus where the active gas was under high pressure. Before the investigation of the electron emission was started, the insulation of the electrometer apparatus and of the quartz cylinder surrounding the platinum electrode were tested each time. The whole vacuum apparatus was also tested for residual gases. Usually, in the vacuum it was impossible to notice any emission ( $1.0 \times 10^{-14}$  amp. would have been detected). The gases  $\text{NOCl}$  and  $\text{Cl}_2$ , however, are exceptions to this. With these, in spite of prolonged pumping out of the system for several hours, after the testing of the gas was finished, it was still possible to observe some little residual emission of the order of  $10^{-13}$ – $10^{-12}$  amp. The platinum cylinder, which was kept at about  $100^\circ \text{C}$ ., served as an electrode. When the system was taken to pieces for the renewal of the alloy the reaction chamber was cleaned and the electrode dismantled. The electrode was set up again only after scrupulous cleaning with acids and heating in a vacuum at  $1000^\circ \text{C}$ . for several hours. The composition of the alloy was the same in all the experiments: 77% K and 23% Na, which corresponds to the formula  $\text{NaK}_3$ . The liquefied gases  $\text{NOCl}$ ,  $\text{Cl}_2$ , and  $\text{COS}$  were supplied by Schering-Kahlbaum, A.G., Berlin; liquid  $\text{Br}_2$  and solid  $\text{I}_2$  by Griffin & Tatlock, Ltd., London; liquefied  $\text{N}_2\text{O}$  by the British Oxygen Co., Ltd., Wembley; for the vapour of  $\text{H}_2\text{O}$  distilled water was used.

#### § 4. *Results of the Experiments.*

We shall first submit some general information about the contact potential difference and reverse current observed in these experiments.

Variation in contact potential difference with change in pressure was observed with all gases which gave the electron emission, and the character of this variation was the same for all of them. When the pressure was increased the contact potential difference increased also and then continued to grow slowly in spite of the fact that the pressure remained constant. Usually, as, for example, with  $\text{COCl}_2$ , or with all other gases at very low pressures, this change in contact potential difference during 2 hours was of the order of 0.03 volt. Since for each characteristic curve,  $i = f(V_1)$ , three consecutive series of observations were usually made, with about 20 measurements in each series, it was easy to introduce a correction for such a change in the contact potential difference during the experiment. At very low pressures of the order of  $10^{-6}$  mm., the observed increase in contact potential difference with time was least for  $\text{Cl}_2$  and perhaps for  $\text{NOCl}$ . However, for these gases the change in contact potential difference with pressure was especially noticeable, and, beginning with pressures of about  $10^{-5}$  mm., the change in contact potential difference with time became comparatively large. For example, with  $\text{NOCl}$  at  $p = 2.4 \times 10^{-4}$  mm. the ratio  $i/i_0$  at +2 volts was originally 0.140 and during 2 hours dropped down to 0.036, which corresponds to an increase in contact potential difference of 0.3 volt. The general character of the changes in contact potential difference indicates that the active gas alters the surface conditions of the platinum cylinder and that these variations are largely reversible. As pointed out in Parts I and II this is probably due to adsorption of gas on the surface of the cylinder. It is interesting to note that in one experiment (after testing  $\text{NOCl}$  and  $\text{Cl}_2$ ) when taking the apparatus to pieces for cleaning, a light grey layer in the form of a ring was noticed on the inner surface of the cylinder. It looks as if in some circumstances the molecules of the alloy may also reach the surface of the cylinder.

Another accompanying phenomenon is a reverse current. In a sufficiently high electron-retarding field the current changes its direction and with further increase in the intensity of the field assumes some constant value. For almost all the gases investigated this value was about 1/10000 of the value of the electron saturation current. It is interesting to note that no such reverse current was observed at all with  $\text{COCl}_2$ . In Table I data for various gases are given, from which it may be seen how the ratio of the reverse current  $i_0'$  to the electron saturation current  $i_0$  changes with pressure. Some increase of

the ratio with pressure is clearly noticeable for the gases NOCl and Cl<sub>2</sub>, while for the gases Br<sub>2</sub> and I<sub>2</sub> the ratio shows no such increase. COS stands out with a considerably higher ratio, but it is necessary to say here that owing to very small currents the ratio could be measured only very roughly. With other gases in view of the constancy of  $i_0'$  during the whole experiment the presence of such a current did not interfere with the accuracy of the measurements of the very small values  $i/i_0$ , which are used for the determination of the shape of the curve for energy distribution in the region of high energies of electrons. The true value of the electron current at a certain retarding field was determined by adding the magnitude  $i_0'$  to the measured value of  $i$ . This is equivalent to assuming that the reverse current had already become saturated with a retarding field at which the direct current was still relatively large.

Table I.

Pressure.	Gas.											
	NOCl.			Cl <sub>2</sub> .			Br <sub>2</sub> .			I <sub>2</sub> .		COS.
$p$ in $1 \times 10^{-6}$ mm.	0.3	2.0	24	0.3	0.3	41	0.5	1.3	3.7	0.5	1.3	150
$i_0'/i_0$ in $10^{-4}$	0.5	1.8	2.2	0.7	0.5	8.2	16	14	11	11	11	ca. 500
$i_0$ in $1 \times 10^{-8}$ amp.	1.7	20	31.5	10.1	9.2	28	1.0	2.3	14.8	0.16	0.41	0.0014

Now we come to the systematic description of the experimental results.

1. NOCl Gas (*Nitrosyl Chloride*).—In view of the considerable change in contact potential difference with time at higher pressures the characteristic curves,  $i = f(V_1)$ , were obtained only at low pressures (this refers also to Cl<sub>2</sub>). One of the right-hand curves in fig. 1, experimental points marked O, corresponds to the pressure  $p = 3 \times 10^{-6}$  mm., time of drop  $T = 54$  seconds, exposure  $dt = 25$  seconds, saturation current  $i_0 = 17 \times 10^{-10}$  amp. The curve is arbitrarily shifted by 1.84 volts to the right along the volt axis. Another characteristic curve for which the experimental points, marked ●, are shifted to the right by 2.56 volts, corresponds to the same pressure  $p = 3 \times 10^{-6}$  mm., but  $T = 51$  seconds,  $dt = 25$  seconds,  $i_0 = 17 \times 10^{-10}$  amp. The two right-hand curves in fig. 2 are the  $\log i/i_0$  vs.  $V_1$  plots with the experimental points marked to correspond with fig. 1. NOCl (O) is shifted to the right by 1.64 volts and NOCl (●) by 2.46 volts. The true zeros ( $V = 0$ ) are indicated by arrows. Fig. 3 represents the function  $i_0 = f(p)$  at  $V_1 = 4$  volts connecting the saturation current  $i_0$  and the gas pressure  $p$ . In the upper part of the diagram the relation between  $\log i_0$  and  $\log p$  is given for that part of



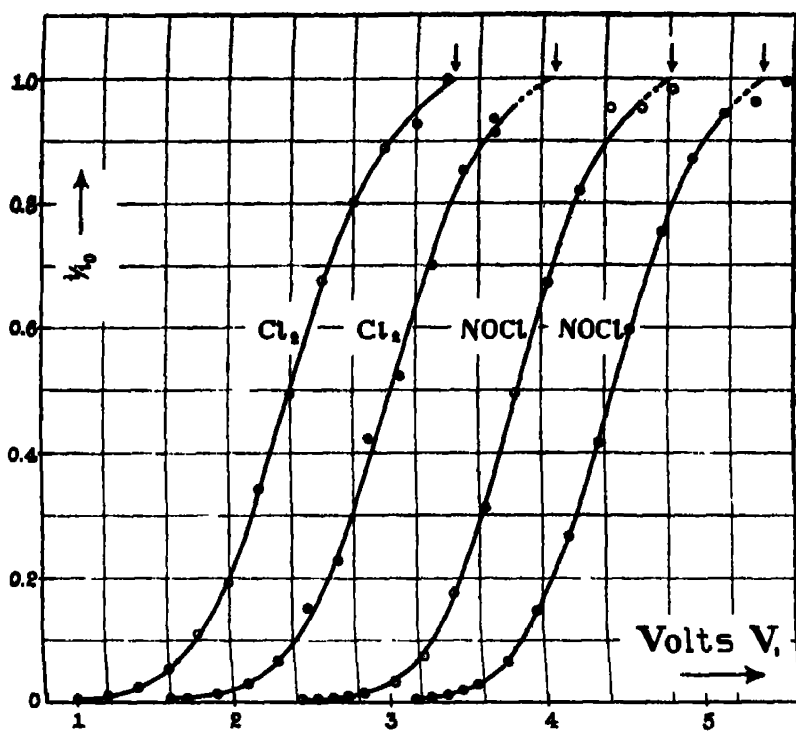


FIG. 1.

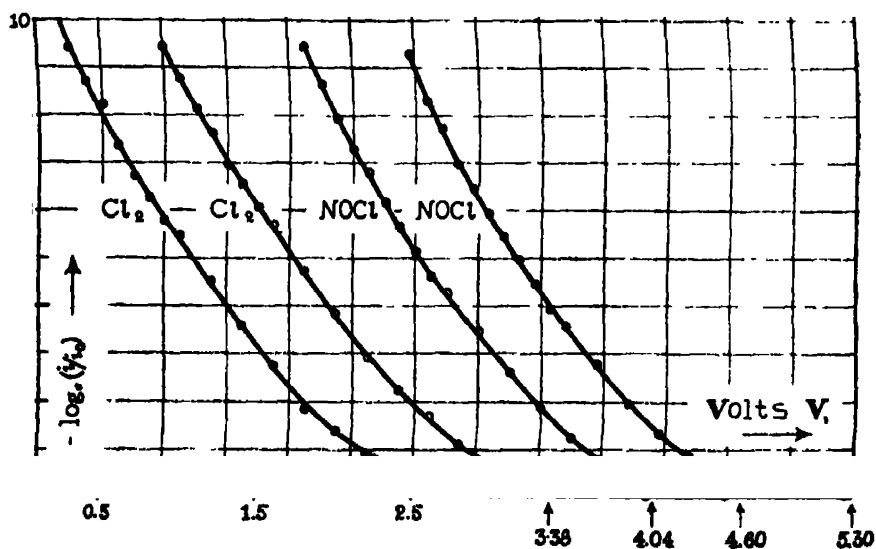


FIG. 2.

the curve  $i_0 = f(p)$  in which the decrease of the emission with increase of pressure was observed. The diagram shows that the ratio between  $i_0$  and  $p$  can be expressed by the formula

$$(i_0)^\alpha p = \text{constant},$$

where  $\alpha = 1.82$ . In this experiment  $T = 45$  seconds,  $dt = 25$  seconds,  $(i_0)_{\text{max.}} = 55 \times 10^{-9}$  amp., which is about four times as much as for  $\text{COCl}_2$ , conditions with regard to change of the drop surface of the alloy with time being the same.

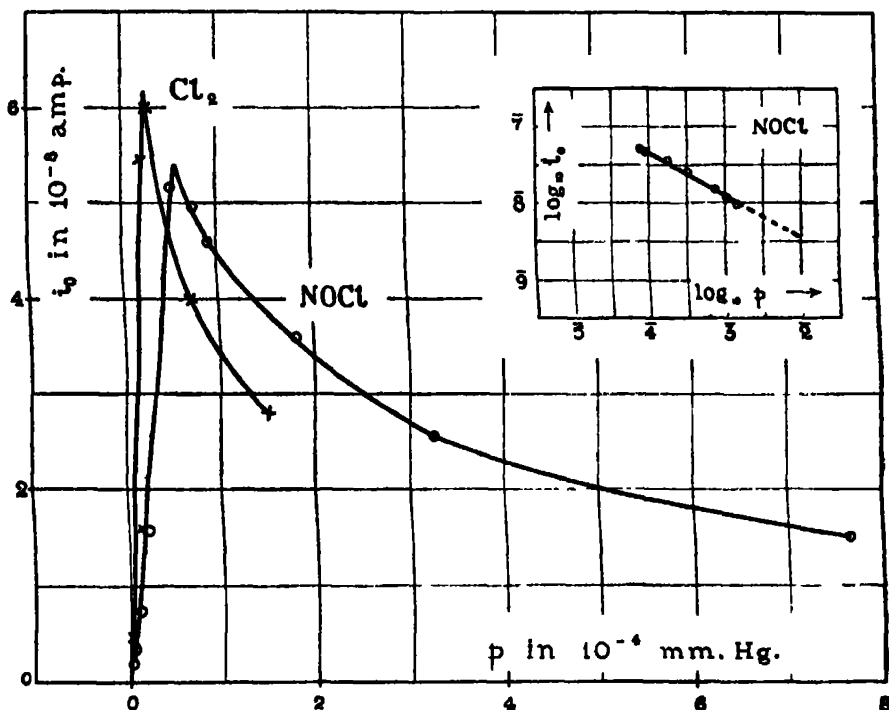


FIG. 3.

2.  $\text{Cl}_2$  Gas (Chlorine).—Two characteristic curves were obtained and are shown in the left part of fig. 1. The first, experimental points marked  $\circ$ , corresponds to the pressure  $p = 4 \times 10^{-6}$  mm.,  $T = 35$  seconds,  $dt = 25$  seconds,  $i_0 = 10.1 \times 10^{-9}$  amp. The second curve, points thus  $\bullet$ , is shifted by 0.7 volt to the right and is at approximately the same pressure  $p = 4 \times 10^{-6}$  mm.,  $T = 29$  seconds,  $dt = 25$  seconds,  $i_0 = 9.2 \times 10^{-9}$  amp. The  $\log i/i_0$  vs.  $V_1$  plots for these two experiments are shown in fig. 2. The points are marked in the same way as the corresponding data in fig. 1.  $\text{Cl}_2$  ( $\circ$ ) is shifted by 0.6 volt to

the right from  $V_1 = 0$ . In fig. 3 is given the function  $i_0 = f(p)$  at  $V_1 = 5$  volts, the points for  $\text{Cl}_2$  being marked with crosses;  $T = 45$  seconds,  $dt = 25$  seconds. The curve for  $\text{NOCl}$ , as has been stated already, is given in the same diagram. It is necessary to point out that it is possible only approximately to compare these two curves as the absolute value of the pressures is not known accurately (the calibration of the Pirani gauge was carried out only for  $\text{N}_2$ ). From the comparison of the curves it is clear that the yield of electrons for  $\text{Cl}_2$  and  $\text{NOCl}$  is not very different. In fig. 4 on the right side the same curve for  $\text{Cl}_2$  is shown once more separately. The curves on the left show how the electron

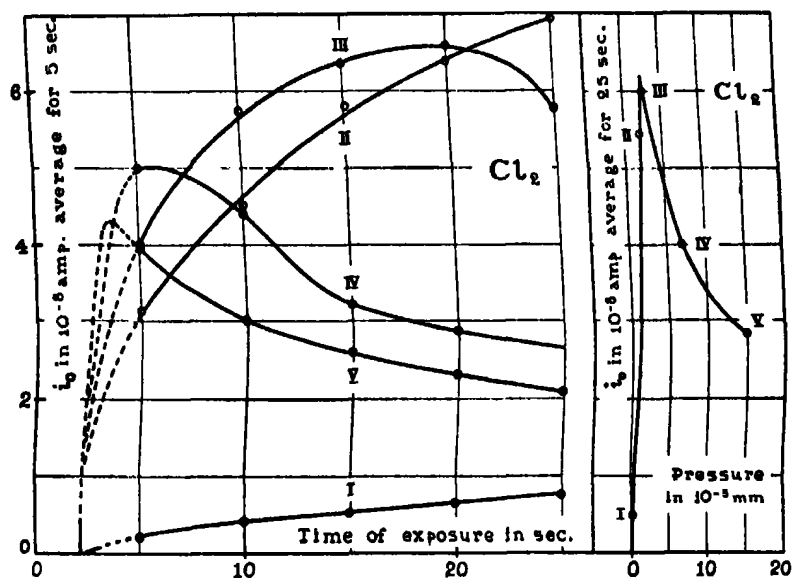


FIG. 4.

current (average for 5 seconds) is connected with the time of exposure  $t$  for five different pressures:  $3.6 \times 10^{-6}$  mm.,  $2.0 \times 10^{-5}$  mm.,  $2.7 \times 10^{-4}$  mm.,  $7.0 \times 10^{-5}$  mm., and  $15 \times 10^{-5}$  mm. The time of drop in all these experiments was about 45 seconds. The five points shown on the curve  $i_0 = f(p)$  correspond to five curves in the left part of fig. 4 and they represent the average current for 25 seconds. From fig. 4 it is possible to see quite clearly the hindering influence of the reaction products on the electron emission. From the curve I we can reach the conclusion that the drop surface increases steadily with the time; other curves show that at a sufficiently high pressure the emission decreases in spite of the increase in the drop surface. Moreover, as an approximate calculation shows, this hindering influence becomes very noticeable when

a more or less complete *monomolecular* layer of the reaction products is formed on the drop surface ( $p \approx 2.5 \times 10^{-6}$  mm.).

3. *HCl Gas (Hydrogen Chloride).*—The results of the measurements of the emission at  $V_1 = 4$  volts (some measurements were made also at  $V_1 = 6$  volts but no striking change was noticed) for various pressures are given in Table II.

Table II.

Pressure of HCl in $1 \times 10^{-6}$ mm. . . . .	2.2	3.6	11	23
Current $i_0$ in $1 \times 10^{-14}$ amp. . . . .	33	74	16	16

The largest current observed was only  $i_0 = 74 \times 10^{-14}$  amp., at pressure  $p = 3.5 \times 10^{-6}$  mm. With increase in pressure the current dropped sharply. Owing to the very small current the determination of the characteristic curves was not carried out. At sufficiently high pressures (about  $10^{-4}$  mm.) the change in the glitter of the drops was clearly noticeable. On the surface of the alloy in the lower bottle, where the alloy drops were falling down after the reaction, the formation of the reaction products was also noted.

4. *Br<sub>2</sub> Gas (Bromine).*—The first of the left-hand curves in fig. 5, experimental points marked O, corresponds to  $p = 5 \times 10^{-6}$  mm.,  $T = 21$  seconds,  $dt = 20$  seconds,  $i_0 = 1.03 \times 10^{-9}$  amp. The curve, with experimental points marked ●, was taken under the following conditions:  $p = 1.3 \times 10^{-6}$  mm.,  $T = 23$  seconds,  $dt = 20$  seconds,  $i_0 = 2.32 \times 10^{-9}$  amp. This curve is shifted arbitrarily 0.6 volt to the right from  $V_1 = 0$ . The corresponding  $\log i/i_0$  vs.  $V_1$  plots are shown in fig. 6. Br<sub>2</sub> (●) is shifted by 0.6 volt to the right from  $V_1 = 0$ . In carrying out the measurements of the pressure of bromine, and especially of iodine, by the Pirani gauge many abnormalities were noticed. Thus the readings of the gauge were not consistent, and for iodine, as a rule, it was difficult to notice the changes in the resistance of the Pirani gauge at low pressures. For these two gases, therefore, the pressure  $p$  was ascertained by calculation on the basis of the following data. The saturation vapour pressure at the temperature concerned was taken from the tables. This is equal to the pressure  $P$  of the vapour of Br<sub>2</sub>, or I<sub>2</sub>, in the large sphere (see fig. 1, Part I). The value of  $p$  was deduced from the ratio  $p/P$  got from the measurements with COCl<sub>2</sub>, introducing a corresponding correction for the difference of molecular weights of COCl<sub>2</sub> and Br<sub>2</sub>, or I<sub>2</sub>. With regard to the function  $i_0 = f(p)$  for Br<sub>2</sub> and I<sub>2</sub>, it was decided to limit the observations to the region of low pressures only, as these gases act very badly on the grease of the vacuum taps. This is of special importance for the high vacuum part

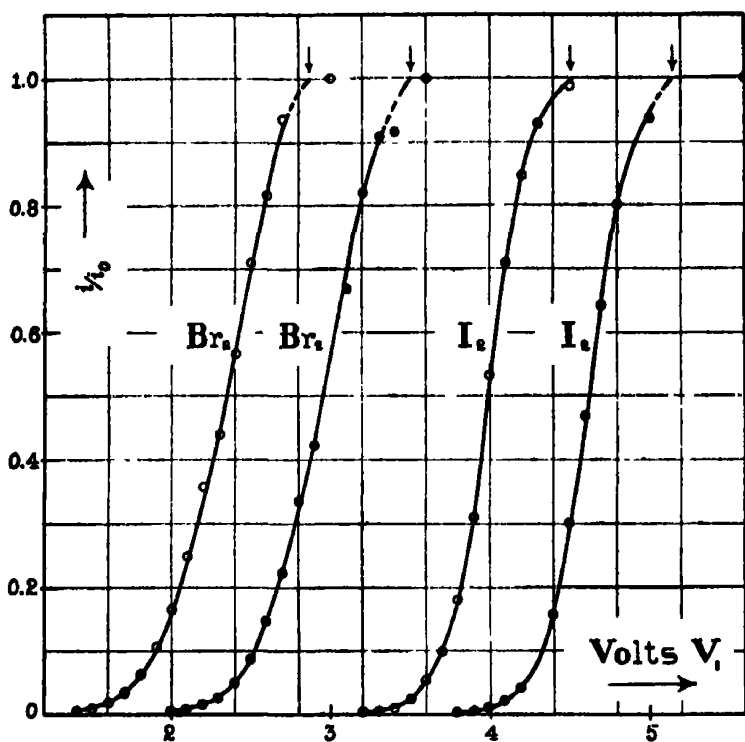


FIG. 5.

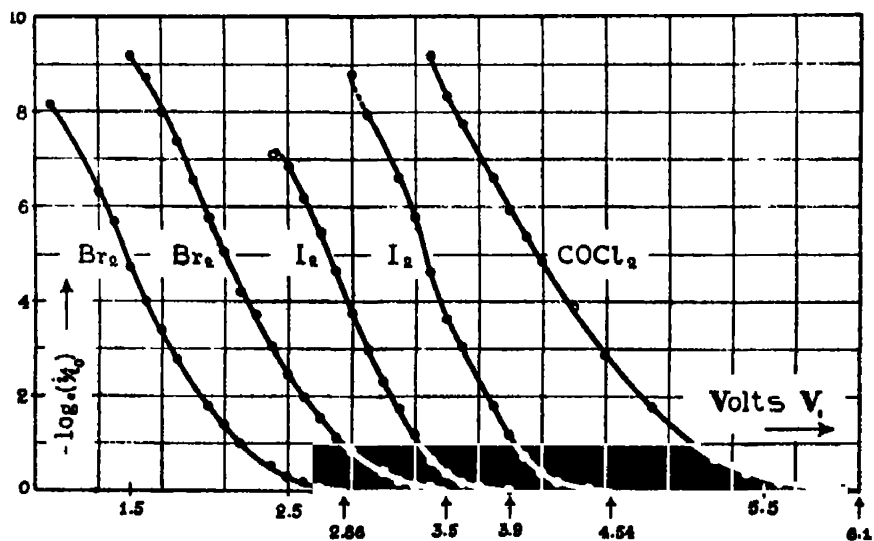


FIG. 6.

of the glass system where the changing of grease on ground joints is very complicated. In Table III, column 3, the ratios  $i_0/p$  are given at the different pressures. We see that the yield  $i_0/p$  of electrons for  $\text{Br}_2$  increases at the higher pressures ( $< 10^{-4}$  mm.) in a similar way to that observed with  $\text{NOCl}$  or  $\text{Cl}_2$ , but this increase may have little significance owing to the uncertainty in the pressures.

Table III.

NOCl.		$\text{Cl}_2$ .		$\text{Br}_2$ .		$\text{I}_2$ .	
T = 45 seconds.		T = 45 seconds.		T = 23 seconds.		T = 22 seconds.	
$p$ in $10^{-5}$ mm.	$i_0/p$ in $10^{-4}$ amp./mm.	$p$ in $10^{-5}$ mm.	$i_0/p$ in $10^{-4}$ amp./mm.	$p$ in $10^{-5}$ mm.	$i_0/p$ in $10^{-4}$ amp./mm.	$p$ in $10^{-5}$ mm.	$i_0/p$ in $10^{-4}$ amp./mm.
0.3	7.3	0.2	14	0.3	1.8	0.04	0.29
0.56	6.3	0.35	14	0.65	1.9	0.08	0.31
1.1	6.8	1.5	11	1.3	1.8	0.17	0.31
2.2	7.2	2.0	27	2.5	3.0	0.38	0.35
5.2	9.9	2.7	22	3.7	4.0	0.67	0.32
				6.1	5.2	0.97	0.30
						1.6	0.29

5.  $\text{I}_2$  Gas (Iodine).—Two characteristic curves for  $\text{I}_2$  are shown on the right of fig. 5. Experimental points marked  $\bigcirc$  correspond to  $p = 5.1 \times 10^{-6}$  mm.,  $T = 22$  seconds,  $dt = 20$  seconds,  $i_0 = 1.57 \times 10^{-10}$  amp. and points marked  $\bullet$  refer to  $p = 1.3 \times 10^{-5}$  mm.,  $T = 22$  seconds,  $dt = 20$  seconds, and  $i_0 = 4.08 \times 10^{-10}$  amp.  $\text{I}_2 (\bigcirc)$  and  $\text{I}_2 (\bullet)$  are shifted to the right from  $V_1 = 0$  by 2 and 2.6 volts respectively. The corresponding  $\log i/i_0$  vs.  $V_1$  plots are shown in fig. 6. In the same figure one of the curves for  $\text{COCl}_2$ , obtained in the previous researches, is drawn for comparison,  $\text{I}_2 (\bigcirc)$ ,  $\text{I}_2 (\bullet)$ , and  $\text{COCl}_2 (\bigcirc)$  are arbitrarily shifted to the right from  $V_1 = 0$  by 1.4, 2, and 4.2 volts respectively. It will be seen from fig. 6 that the energy distributions with  $\text{Br}_2$  and  $\text{I}_2$  are markedly different from that with  $\text{COCl}_2$ . In Table III, column 4, are given the values of  $i_0/p$  at the different pressures. The yield of electrons is roughly one-sixth of that for  $\text{Br}_2$ . It will be noticed that in each of the gases  $\text{NOCl}$ ,  $\text{Cl}_2$ ,  $\text{Br}_2$ , and  $\text{I}_2$  the yield  $i_0/p$  is independent of  $p$  for values of  $p$  less than about  $1.5 \times 10^{-5}$  mm.

6.  $\text{H}_2\text{O}$  Gas (Water Vapour).—With this gas the first observations of the emission were altogether different from all the later ones, and therefore we give below the description of the whole experiment. One c.c. of distilled water

was used for obtaining the water vapour. The tube containing the water was kept at  $-78^{\circ}\text{C}$ . (solid  $\text{CO}_2$  in acetone) during the pumping out of the whole system. After having finished up all the preliminary pumping out, the Dewar vessel with the freezing mixture was put aside in order to enable the generating tube to acquire room temperature. During the melting of the water bubbles of gas were visible in considerable quantity, and the first measurements showed that the emission was as high as  $3 \times 10^{-11}$  amp. at about  $p = 5 \times 10^{-5}$  mm. Then the second pumping out was undertaken. When the water melted this time no visible signs of gas evolution were noticed. The testing of the emission also gave quite different results. Thus we see from Table IV that the emission did not exceed  $25 \times 10^{-14}$  amp. even at such a high pressure as  $9 \times 10^{-4}$  mm.

Table IV.

Pressure of $\text{H}_2\text{O}$ $p$ in $1 \times 10^{-6}$ mm.	0.4	0.6	0.9	2.3	3.6	6.3	12	23	40	90
$i_0$ in $1 \times 10^{-14}$ amp.	0.0	2	8	0.0	9	9	9	12	20	25

That the reaction took place was evident from the slight change in the glitter of the drops and by a quite noticeable white deposit on the alloy surface in the lower bottle.

7.  $\text{N}_2\text{O}$  Gas (*Nitrous Oxide*).—With regard to this gas Table V shows that no emission was noticed up to pressures of the order of  $5 \times 10^{-4}$  mm. No sign was in evidence that any reaction generally took place.

Table V.

Pressure of $\text{N}_2\text{O}$ $p$ in $1 \times 10^{-5}$ mm.	0.7	1.6	4.9	8.6	50
$i$ in $1 \times 10^{-14}$ amp.	0.0	0.0	1.0	0.0	0.0

8.  $\text{COS}$  Gas (*Carbon Oxysulphide*).—The negative emission observed with this gas represents in general a very small magnitude. Fig. 7 shows the change of the emission with pressure. The curve was determined under the experimental conditions:  $V_1 = 4$  volts,  $T = 27$  seconds,  $dt = 20$  seconds. At a sufficiently high pressure (about  $10^{-3}$  mm.) the emission reaches a saturation value which is equal to  $140 \times 10^{-14}$  amp. In fig. 8 the characteristic curve corresponds to  $p = 1.5 \times 10^{-3}$  mm.,  $T = 27$  seconds,  $dt = 20$  seconds,  $i_0 = 140 \times 10^{-14}$  amp. The error of measurement was approximately  $7 \times 10^{-14}$  amp. which made it impossible to determine values of  $i/i_0$  less than 0.05. The exceptionally low accuracy of this particular experiment is probably

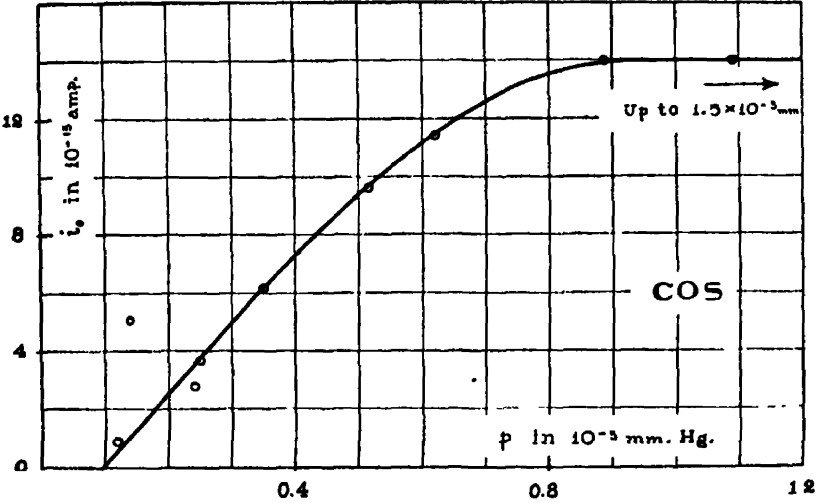


FIG. 7.

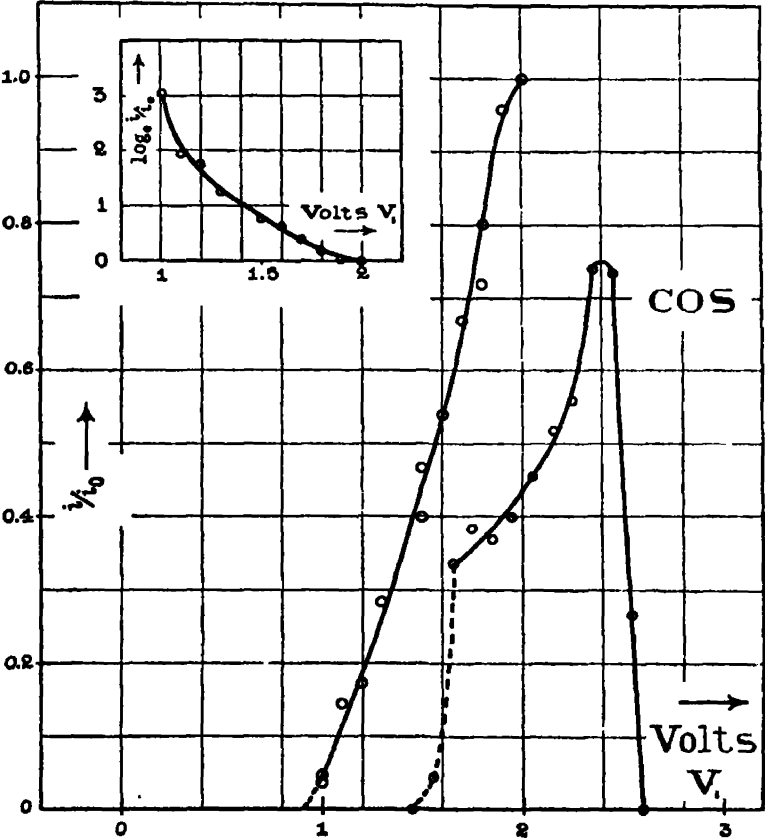


FIG. 8.



due to the high gas pressure which had to be used in order to get measurable currents. In the upper part of fig. 8 the  $\log i/i_0$  vs.  $V_1$  plot is given and it corresponds to the characteristic curve. For this gas the change in glitter of drops was very noticeable, which indicates that the reaction took place.

### § 5.

The curves showing the energy distribution among the electrons emitted by  $\text{COCl}_2$ ,  $\text{NOCl}$ , and  $\text{Cl}_2$  are shown in fig. 9. They are constructed in the same

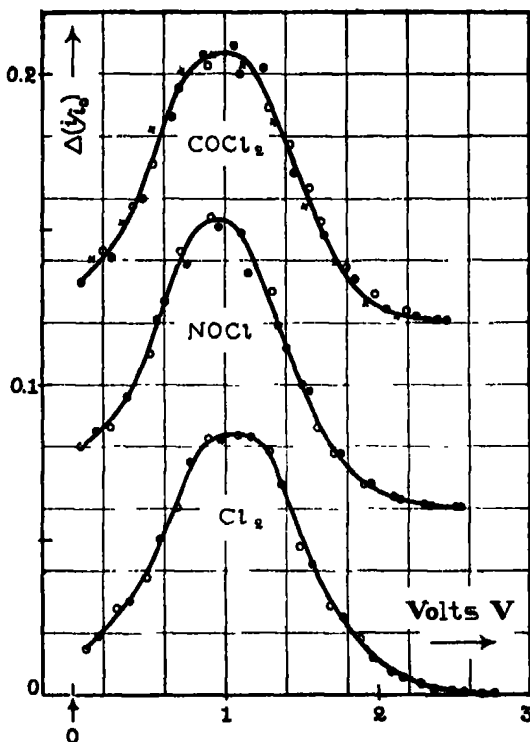


FIG. 9.

way as those of fig. 4 of Part II. In fact, the curve for  $\text{COCl}_2$  in fig. 9 is just a repetition of fig. 4 of Part II. The abscissæ are volts and the ordinates the values of  $\Delta(i/i_0)$  for each increment  $\Delta V$  of 0.1 volt. They are therefore proportional to the fraction of the total number of electrons which have kinetic energies within the range  $eV$  to  $e(V + \Delta V)$ . In the graphs for  $\text{NOCl}$  and  $\text{Cl}_2$  the points are marked in the same way as the corresponding data in figs. 1 and 2 from which they are derived. The curves for  $\text{NOCl}$  and  $\text{COCl}_2$  in fig. 7 are displaced vertically upwards by the arbitrary amounts: 0.05 and 0.1 unit respectively.

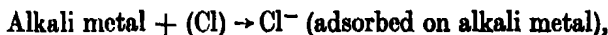
The energy distribution curves for these three gases are remarkably alike. Even the "tails" at the high energy end of the spectrum are all very similar. The maximum values of  $\Delta(i/i_0)$  are also almost identical. Among the differences between them which are just discernible we may mention (1) the curve for  $\text{Cl}_2$  is a little wider and that for  $\text{NOCl}$  a little narrower than that for  $\text{COCl}_2$ ; (2) the curve for  $\text{Cl}_2$  extends a little farther to the right than the others, showing that the electrons emitted by this gas have somewhat higher energies. We saw in Part II that the maximum value  $E_m$  of the energy of the electrons emitted with  $\text{COCl}_2$  was estimated to be  $2.5_3 \pm 0.1$  volts. An estimate carried out in a precisely similar manner gave a value of  $E_m$  for  $\text{NOCl}$  equal to  $2.6_0 \pm 0.1$  volts, and for  $\text{Cl}_2$   $2.8_0 \pm 0.1$  volts. As pointed out in Part II it is difficult to determine  $E_m$  very precisely, partly on account of the tails and sometimes also on account of the difficulty of locating the true zero or saturation point. However, for the three gases  $\text{COCl}_2$ ,  $\text{NOCl}$ , and  $\text{Cl}_2$ , which all give large and steady emissions and for which the distribution curves are so very much alike, it is very unlikely that the differences of the values of  $E_m$  determined in this way will disagree with the true differences by more than 0.2 volt. This conclusion is of importance for the interpretation of the character of the chemical reactions involved, as we shall see in § 6.

Fig. 10 shows the energy distribution curves for  $\text{Br}_2$  and  $\text{I}_2$  and also repeats the  $\text{Cl}_2$  curve of fig. 9 for comparison. The points are marked in the same manner as the corresponding data in figs. 1 and 5 on which they are based. The curves for  $\text{Br}_2$  and  $\text{Cl}_2$  are arbitrarily shifted to the right from  $V = 0$  by 0.7 and 1.1 volts respectively. The true zeros are indicated by arrows. In contrast to the curves for  $\text{COCl}_2$ ,  $\text{NOCl}$ , and  $\text{Cl}_2$  the curves for  $\text{Cl}_2$ ,  $\text{Br}_2$ , and  $\text{I}_2$  are more conspicuous for their differences than for their points of resemblance. The curves become progressively narrower as we proceed from  $\text{Cl}_2$  through  $\text{Br}_2$  to  $\text{I}_2$ , which means that there is a corresponding diminution in the average energy of the electrons emitted (as well as of the maximum energy). We also notice another difference. The curve for  $\text{Cl}_2$ , like those for  $\text{COCl}_2$  and  $\text{NOCl}$ , is nearly, although not quite, symmetrical about a vertical line passing through its maximum. The curves for  $\text{Br}_2$  and  $\text{I}_2$  are much more asymmetrical about such a line. The value of  $E_m$  for the main part of the emission from  $\text{Br}_2$  is close to  $1.7_0$  and that for the main part of the emission from  $\text{I}_2$  to  $1.5_0$  volts.

There is, however, some evidence of the existence in the reactions with these gases of a small emission having higher energy than the main part. This can be seen if we compare the  $\log i/i_0$  vs.  $V_1$  plots of  $\text{Br}_2$  and  $\text{I}_2$ , fig. 6, with those for  $\text{COCl}_2$ , fig. 6,  $\text{NOCl}$  and  $\text{Cl}_2$ , fig. 2. The curves for  $\text{COCl}_2$ ,  $\text{NOCl}$ , and  $\text{Cl}_2$

are obtained in this Section, and in the next one, are also of some interest in connection with the problem of the kinetics of heterogeneous chemical reactions. In this respect, the chemical electron emission experiments may be regarded as a new method of investigating the kinetics of heterogeneous chemical reactions.

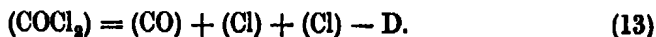
In Part II, § 6, we gave reasons for believing that in a reaction such as that between  $\text{COCl}_2$  and the liquid alloy of sodium and potassium the emission of electrons took place at the instant of formation of a polar bond between a chlorine atom of the  $\text{COCl}_2$  and an alkali atom of the alloy, or during an interval subsequent to this by not more than about  $10^{-13}$  seconds. We assumed that the reaction occurred with a normal average collision diameter  $r_1$ , which we put approximately equal to the sum of the radii of the chlorine and alkali atoms. On this basis we derived, for the energy  $E_e$  available at the instant of the electronic rearrangement in the reaction



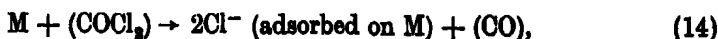
the following expression :—

$$E_e = \frac{e^2}{r_1} + \frac{e^2}{4} \frac{1}{x_0 + r_{\text{Cl}}} + A_{\text{Cl}} - \phi_{\text{K,Na}}. \quad (12)$$

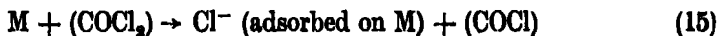
In (12)  $e$  is the electronic charge,  $r_{\text{Cl}}$  is the radius and  $A_{\text{Cl}}$  the electron affinity of the chlorine atom,  $\phi_{\text{K,Na}}$  is the work function of the alloy and is  $= \frac{1}{4} \frac{e^2}{x_0}$ , and  $\frac{e^2}{4} \frac{1}{x_0 + r_{\text{Cl}}}$  is the negative potential energy of the negative Cl ion when at  $r_1$  due to the presence of the metal surface and calculated from the image force. We then showed that equation (12) was in agreement with the observed maximum energy  $E_m$  of the emitted electrons if we made the further assumption that with  $\text{COCl}_2$  both chlorine atoms simultaneously entered into the reaction. That is to say, we assumed that the maximum electron energy  $E_m$  would be equal to the quantity obtained by subtracting from the maximum available energy  $E_e$  given by (12) the sum of the quantities  $\phi_{\text{K,Na}}$ , which is the work an electron has to do to get out of the metal, and  $\frac{1}{2}D$  the energy of dissociation of  $(\text{COCl}_2)$  according to the reaction



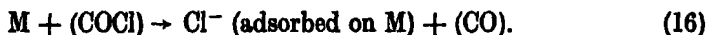
This is equivalent to the assumption that the reaction between  $(\text{COCl}_2)$  and the alloy goes to a considerable extent in one stage



where M is an abbreviation for the alkali metal. Obviously, however, there is another possible mechanism for the reaction, namely, the two-stage



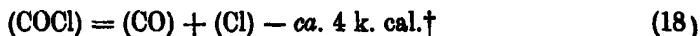
and



Instead of subtracting  $\frac{1}{2}D$  belonging to the reaction (13) we should, if the reaction goes like (15), have to subtract the heat of the reaction



and if it goes like (16), that of the reaction



We disregard the possibility<sup>‡</sup> that the (CO) radicle will enter into the reaction in view of its high dissociation energy:  $(\text{CO}) = (\text{C}) + (\text{O}) - ca. 250 \text{ k. cal.}^\S$

If the reaction is in two stages we expect the first stage to predominate because the radicle, such as COCl, which is liberated will in general fly off into the surrounding space and have little chance of reacting with the metal drop (compare the next § 7).

In Part II, § 6, we found  $E_0$  of equation (12) to be 7.0<sub>8</sub> volts; so that

$$E_0 - \phi_{K,Na} = 4.5_8 \text{ volts.} \quad (19)$$

On subtracting  $\frac{1}{2}D = 42 \text{ k. cal.} = 1.83 \text{ volts}$  (the value according to (13)) we obtained 2.7<sub>8</sub> volts whereas the measured maximum electron energy was between 2.5<sub>2</sub> and 3.0<sub>5</sub> volts. If, however, the reaction goes according to (15) we must subtract the volt equivalent of 80 k. cal., or 3.48 volts instead of 1.83, giving 1.10 volts. If it goes according to (16) we must subtract  $4 \div 23.07 = 0.18$  volt which gives 4.40 volts. Neither of these values shows any agreement with the observed data; consequently unless the view we have taken as to the nature of the omission process is wrong the mechanism responsible

\* Bodenstein, Lenher, and Wagner, 'Z. phys. Chem.,' B, vol. 3, p. 459 (1929); Herzberg, 'Ergebn. exact. Naturwiss.,' vol. 10, p. 278 (1931).

† Bodenstein, Lenher, and Wagner, *loc. cit.*; Herzberg, *loc. cit.*

‡ This conclusion is also supported by the fact that free radicles, such as CO, NO, etc., have in general, in contrast to the free halogen atoms, an appreciable inertia, i.e., the reactions between the radicles and the alloy probably require an appreciable (true) activation energy.

§ Unless other references are given all thermochemical data in the paper are taken from Landolt-Bornstein, 'Tabellen' (1924-31).

for the electron emission with  $\text{COCl}_2$  must be largely one stage and according to (14). It is not, however, excluded on this evidence that with  $\text{COCl}_2$  the two-stage reaction (15) and (16) may not be occurring to some extent. The amount in the second stage, going according to (16), might be too small to detect and there may be a group of slow electrons coming from the first stage (15) superposed on the main group.

Further light on these problems is obtained when we consider the other reactions involving chlorine. For  $\text{NOCl}$  the only possible course of the reaction is



since we consider that  $(\text{NO})$  will not take part in the reaction on account of its high dissociation energy:  $(\text{NO}) = (\text{N}) + (\text{O}) - \alpha. 150 \text{ k. cal.}$  There is no reason why the average collision diameter  $r_1$  should be appreciably different for the reaction with  $\text{NOCl}$  from its value for the reaction with  $\text{COCl}_2$ . At any rate we make the assumption that the distance  $r_1$  is, to the degree of approximation of the calculation, the same for all reactions involving the chlorine atoms and see where it leads us. The other quantities in equation (12) are the same for all reactions with the alloy which involve chlorine. If the reaction with  $\text{NOCl}$  goes according to (20) then from  $E_e - \phi_{\text{K,Na}} = 4.5_8 \text{ volts}$  we have to subtract  $1.65 \text{ volts}$  ( $= 38 \text{ k. cal.}$ ) since

$$(\text{NOCl}) = (\text{NO}) + (\text{Cl}) - 38 \text{ k. cal.}^* \quad (21)$$

This gives for the maximum kinetic energy for the electrons  $2.9_0 \text{ volts}$  in satisfactory agreement with the experimental maximum which lies between the limits  $2.6_0$  and  $\alpha. 3.1 \text{ volts}$ .

Since for  $\text{NOCl}$  there is only one possible mechanism for the reaction this result determines the value of  $r_1$  and proves that the reaction with  $\text{COCl}_2$  goes to a great extent in one stage, unless  $r_1$  has considerably different values in the reactions with  $\text{NOCl}$  and with  $\text{COCl}_2$ . But it is easy to show that  $r_1$  for  $\text{COCl}_2$  cannot be considerably different from  $r_1$  for  $\text{NOCl}$ . In fact, as in a two-stage reaction the first stage predominates this means that  $r_1$  for  $\text{COCl}_2$  must be less than  $r_1$  for  $\text{NOCl}$  because the energy required for reaction (17) is larger than that for (21). But the decrease in  $r_1$  is effected through  $r_{\text{Cl}}$  since  $r_1 = r_{\text{K}} + r_{\text{Cl}}$  and we can put  $r_{\text{K}}$ , the radius of the potassium metal

\* This is based on  $(\text{NO}) + \frac{1}{2}(\text{Cl}_2) = (\text{NOCl}) + \alpha. 8.5 \text{ k. cal.}$  Cf. Trautz and Wachenheim 'Z. anorg. Chem.,' vol. 97, p. 241 (1916) and Trautz and Schluster (*ibid.*, vol. 136, p. 1 (1924)).

atom, to be approximately constant. Now, we can solve equation (12) in respect to  $r_{Cl}$  taking into account equation (17). We thus obtain

$$r_{Cl} = 0.37 \times 10^{-8} \text{ cm.},$$

i.e., about one-third of the value for NOCl. This is evidently impossible since we have given for  $r_{Cl}$ , for NOCl, the smallest possible value  $1.07 \times 10^{-8}$  cm. equal to the radius of the Cl atom.

For  $Cl_2$  we have, as for  $COCl_2$ , two possibilities. These are the one-stage reaction



and the two-stage reaction



and



Continuing with our assumption that  $r_1$  has still the same value (it can be shown, in the same way as for NOCl, that  $r_1$  must have approximately the same value), and since

$$(Cl_2) = (Cl) + (Cl) - 58 \text{ k. cal.}, \quad (25)$$

to obtain the maximum kinetic energy of the electrons we shall have to subtract from  $4.5_9$  volts  $1.26$  volts ( $= 29$  k. cal.), if the reaction is according to (22), or  $2.51$  volts ( $= 58$  k. cal.) if it is according to (23), and zero volts if it is according to (24). These give  $3.2_7$  volts for the one-stage reaction (22),  $2.0_1$  volts for the first part of the two-stage reaction (23), and  $4.5_9$  volts for the second part (24). The experimental value of the maximum kinetic energy is between the limits  $2.8_0$  volts and *ca.*  $3.4$  volts. The reaction responsible for the electron emission is clearly one-stage (22), but as with  $COCl_2$  it is not excluded that the two-stage reaction may occur to some extent.

We based the above considerations on the method which required a knowledge of the absolute value of the chemical energy involved in the reactions. We can arrive, however, at the same results about the reaction mechanism and the collision diameter by a more general method which is based on a relative comparison of the reactions. So far as we deal with the same end product the maximum electron energy  $E_m$  is a function of  $r_1$  and  $D$  only

$$E_m = f(r_1, D). \quad (26)$$

Moreover, if  $r_1$  is approximately independent of the kind of reacting gas molecules then we should expect, according to our theory of electron emission, that

$$E_m + D = \text{constant}, \quad (27)$$

provided  $E_m$  is ascribed to the right reaction mechanism, i.e., to the right value of  $D$ . A comparison of the measured maximum electron energy with the values of  $D$  resulting from the different alternative chemical mechanisms is shown in Table VI.

Table VI.

Gas.	$E_m$ volts (observed).	One-stage reaction ( $E_m + D$ ) volts.	First stage of two-stage reaction ( $E_m + D$ ) volts.	Second stage of two-stage reaction ( $E_m + D$ ) volts.
$\text{Cl}_2$ . . . . .	2.8 <sub>0</sub>	4.0 <sub>0</sub>	5.3 <sub>1</sub>	2.8 <sub>0</sub>
$\text{NOCl}$ . . . . .	2.6 <sub>0</sub>	4.2 <sub>1</sub>	4.2 <sub>1</sub>	4.2 <sub>1</sub>
$\text{COCl}_2$ . . . . .	2.5 <sub>0</sub>	4.3 <sub>0</sub>	6.0 <sub>1</sub>	2.7 <sub>1</sub>

We have already pointed out that owing to the occurrence of the "tails" at the high energy end of the electronic spectra it is difficult either to define or to measure the absolute value of this maximum very precisely, but when the curves, including the tails, have very nearly the same shapes, and the maxima are always measured in the same way we can be sure that the relative values cannot be far wrong. In the present considerations it is only the relative values that we are concerned with. It is obvious from Table VI, in which the values ascribed to  $E_m$  have all been determined in identical ways, that it is only the one-stage reaction for these three gases for which the values of  $E_m + D$  satisfy equation (27). Since it is highly improbable that equation (27) would be so accurately satisfied with random values of  $\tau_1$  we may conclude that the collision diameter must be approximately the same for all these gases, and therefore, the reaction mechanism responsible for the electron emission must also be the same and in one stage.

In our considerations of the reactions involving chlorine we have not, so far, included the reaction with  $(\text{HCl})$ , because with  $\text{HCl}$  there is a complication which is not present in the three preceding reactions, namely, the possibility of the adsorption of the hydrogen atoms on the surface of alkali metal. However, we can show that if there is any electron emission resulting from the adsorption of the hydrogen atoms it must be comparatively very small. For the adsorption of the hydrogen atoms we can use equation (12) with approximately the same numerical data except for the electron affinity  $A_{\text{H}}$  of the hydrogen atom. This latter quantity is much smaller than that for halogens.\*

\* Cf. van Arkel and de Boer, "Chemische Bindung als elektrostatische Erscheinung" p. 65 (Leipzig, 1931).

Hylleraas\* devised a method of calculating the electron affinity of hydrogen by means of the quantum mechanics, obtaining the value 16.3 k.cal. Since

$$(\text{HCl}) = (\text{H}) + (\text{Cl}) - 102 \text{ k. cal.} \quad (28)$$

the calculation gives a positive value (about 1.0 volt) for  $E_m$  only for the second stage of the two-stage reaction, which corresponds to the adsorption of the free hydrogen atoms. But, as already pointed out, we should expect the rate of the second stage of a two-stage reaction to be small.

The absorption of the chlorine atoms may give rise to three different groups of electrons. If the reaction with (HCl) goes in one stage and the energy of dissociation is equally distributed between the H and Cl atoms this may give a group of electrons with  $E_m = 2.3_e$  volts. When the Cl atom of the HCl molecule first enters into the reaction this may give another group of electrons with  $E_m = 0.1_e$  volt. Still another group of electrons may result from the adsorption of the free chlorine atoms; this group will have  $E_m = 4.5_e$  volts.

Owing to the smallness of the electron currents we have not been able to measure the maximum electron energy. This fact deprives us of the possibility of establishing the reaction mechanism, and therefore, we cannot consider this experiment as evidence for, or against, the electron emission theory. On the other hand, if our view as to the electron emission mechanism is correct then we may conclude that the reaction with (HCl) cannot be in one stage since the yield is very small compared with the reactions involving  $\text{Cl}_2$ ,  $\text{NOCl}$ , or  $\text{COCl}_2$ . This conclusion makes it necessary that the number of electrons in the group with  $E_m = 0.1_e$  volt must be small. This is in agreement with all the experimental evidence; there is, as will be clear when more reactions have been described, a general correlation between the energy of the emitted electrons and the yield, low yields accompanying low energies and *vice versa*.

The different behaviour of the (HCl) molecules, in respect to the reaction mechanism, may be attributed to the great difference between the electron affinities and between the sizes of the H and Cl atoms. The (HCl) molecule also differs from the homopolar  $\text{Cl}_2$  and  $\text{COCl}_2$  molecules by its strongly developed features of heteropolar binding.

### § 7. Comparison of the Three Halogens, Chlorine, Bromine, and Iodine.

An analysis of the electronic spectra for  $\text{Br}_2$  and  $\text{I}_2$  has revealed (see § 5) the presence of a small group of electrons with much higher average energies than

\* 'Z. Physik,' vol. 60, p. 624 (1930); vol. 63, p. 291 (1930).



the main group. As no such high energy group of electrons has been observed in the one-stage reactions with  $\text{NOCl}$ ,  $\text{COCl}_2$ , or  $\text{Cl}_2$ , this fact suggests that the reactions with  $\text{Br}_2$  and  $\text{I}_2$  proceed in two stages to an appreciable extent. This result is in complete agreement with the results which can be obtained from a comparison of the observed maximum electron energy  $E_m$  with the energy  $E_c$  resulting from the different alternative chemical mechanisms. Such a comparison is shown in Table VII. The values of  $E_c$  were calculated by means of equation (12) which involves an assumption that the collision

Table VII.

Gas.	$E_m$ volts (observed).	One-stage reaction ( $E_c - \phi$ ) volts.	First stage of two-stage reaction ( $E_c - \phi$ ) volts.	Second stage of two-stage reaction ( $E_c - \phi$ ) volts.
$\text{Cl}_2$	2.8 <sub>0</sub>	3.2 <sub>7</sub>	2.0 <sub>1</sub>	4.5 <sub>0</sub>
$\text{Br}_2$	1.7 <sub>0</sub>	2.9 <sub>0</sub>	2.0 <sub>0</sub>	3.9 <sub>0</sub>
$\text{I}_2$ ...	1.6 <sub>0</sub>	2.6 <sub>0</sub>	1.8 <sub>0</sub>	3.3 <sub>0</sub>

diameter  $r_1$  is approximately equal to the sum of the radii of the alkali metal and halogen atoms. For the radii of the bromine and iodine atoms we used the values given by Goldschmidt from X-ray measurements (see Landolt-Bornstein Tabellen), namely,  $r_{\text{Br}} = 1.19 \times 10^{-8}$  cm. and  $r_{\text{I}} = 1.36 \times 10^{-8}$  cm. The values of the electron affinity  $A$  and dissociation energy  $D$ , referring to bromine and iodine, as given by van Arkel and de Boer (*loc. cit.*), are: for bromine,  $A = 3.4 \pm 0.3$  volts and  $D = 1.96$  volts, for iodine,  $A = 3.1 \pm 0.2$  volts and  $D = 1.53$  volts. The measured values of the maximum energy  $E_m$  for  $\text{Br}_2$  and  $\text{I}_2$  correspond, within the limits of uncertainty of the calculation, to the values of  $(E_c - \phi)$  calculated for the first stage of the two-stage mechanism, as can be seen from Table VII. This shows that the mechanism of the reactions with  $\text{Br}_2$  and  $\text{I}_2$  is almost exclusively a two-stage type. Since the rate of a two-stage reaction, corresponding to the second stage, must be generally small the observed small group of electrons with high average energy may either result from this second stage or it may come from the one-stage reaction; but, in any case, the two-stage must predominate over the one-stage reaction.

It follows, thus, from our analysis of the electronic spectra of halogens that the one-stage mechanism ceases to dominate the reaction in passing from  $\text{Cl}_2$  to  $\text{Br}_2$  and  $\text{I}_2$ . This result is in accord with all the experimental evidence

about the halogen molecules, which, we think, is responsible for the one, or the other, type of the reaction mechanism. In fact, on the one hand, the internuclear distance and the size of the atoms are increasing in the direction  $\text{Cl}_2 \rightarrow \text{Br}_2 \rightarrow \text{I}_2$ . On the other hand, the electron affinity and the dissociation energy are diminishing in the same direction  $\text{Cl}_2 \rightarrow \text{Br}_2 \rightarrow \text{I}_2$ . It is also necessary to add that probably in the electronic spectrum of  $\text{Cl}_2$  there is already a small admixture of low energy electrons which comes from the first stage of the two-stage reaction, but which our method of analysis is unable, for reasons indicated in § 2, to separate from the main group. This small group of low energy electrons probably makes itself evident by disturbing slightly the symmetry of the energy distribution curve about its maximum (commonest

The important circumstance which must be also taken into account when the rate of a certain reaction is considered is that the yield of electrons is a very sensitive function of the energy  $E_e$  available in the reaction. As for  $\text{Cl}_2$  and  $\text{COCl}_2$  the chemical energy  $E_c$  corresponding to the first stage of the two-stage reaction is much lower than that referring to the one-stage reaction, the rate of the two-stage reaction for these gases may be comparatively quite large and, by all indications, is of the same order of magnitude as the rate of the one-stage reaction. The general discussion on this subject, which is reserved until the majority of the available data on the chlorine compounds shall have been presented, shows that, as a matter of fact, the ratio  $\frac{\text{rate of two-stage reaction}}{\text{rate of one-stage reaction}}$  is always  $> 1$ .

It is more or less evident from these experiments, as well as those which will be described in the following papers, that in the reactions involving chlorine the probability of the occurrence of a one-stage reaction is approximately of the same order of magnitude as the probability of the occurrence of a two-stage reaction. But, if so, then it is not surprising that the probability that the second atom of the halogen molecule will enter into the reaction is a very sensitive function of the internuclear distance of the halogen molecule, since the chemical forces, which act between neutral atoms, decrease exponentially with the distance. The increase in the internuclear distance  $r_e$  of the halogen molecules in the direction  $\text{Cl}_2 \rightarrow \text{Br}_2 \rightarrow \text{I}_2$  is quite appreciable; this can be seen from the data given by Jevons\* from band spectra measurements: for  $\text{Cl}_2$   $r_e = 1.98 \times 10^{-8}$  cm., for  $\text{Br}_2 = 2.28 \times 10^{-8}$  cm., and for  $\text{I}_2 = 2.66 \times 10^{-8}$  cm.

\* "Report on Band Spectra of Diatomic Molecules," p. 280, Phys. Soc. Lond. (1932).

### § 8. *Discussion of the Results with COS, N<sub>2</sub>O, and H<sub>2</sub>O.*

In this section we propose to discuss, briefly, the reactions which result in the adsorption of atoms of a type considerably different from those previously investigated, namely, the oxygen and sulphur atoms. They are no longer univalent atoms, like halogens, but they have two free valencies. In the heterogeneous reaction with the alkali metal there is, therefore, the possibility of a simultaneous formation of two polar bonds, such as K-O-K. A detailed study of this sort of reaction is much more complicated on account of the fact that the observed yields were generally very small and the number of reactions available for the investigation is very limited because molecules of most of the oxygen, or sulphur, compounds have more or less completely saturated valencies, and therefore at ordinary conditions they are chemically inactive. The data which we have obtained in the investigation of this group of reactions, namely, with COS, N<sub>2</sub>O, H<sub>2</sub>O, H<sub>2</sub>S, and O<sub>3</sub> (the last two gases to be described in a later communication) seems to be insufficient to take any very definite view about the mechanism of the observed negative emission. This is partly due to the uncertainty of the chemical mechanism of adsorption of the O and S atoms and the difficulty of calculating the chemical energy by a method similar to that for halogens. In view of the currents being of an entirely different order of magnitude there even arises some doubt if the carriers of the currents are electrons at all and not heavy ions. Nevertheless, the results obtained in the investigation of this type of reaction are instructive in giving some useful information about the magnitude of the yield and the maximum energy of the negative emission.

We shall first consider the reaction with carbon oxysulphide as the dissociation process of this gas is very simple and the currents obtained, though being very small (maximum  $140 \times 10^{-14}$  amp. or less than  $10^{-4}$  times the amount obtained with Cl<sub>2</sub>), were still larger than those with N<sub>2</sub>O and H<sub>2</sub>O. According to the dissociation process of COS

$$(\text{COS}) = (\text{CO}) + (\text{S}) - 3.1, \text{ volts}, \quad (29)$$

the observed emission, which shows a maximum energy of the order of 1 volt, may only result from the adsorption of the S atoms. This experiment seems to indicate that at the first moment of collision of the S atom with the alloy the S atom reacts only with one atom of the alkali metal. In fact, under such an assumption the chemical energy  $E_c$  may be estimated by means of equation (12) and the calculation gives for  $E_m = E_c - \phi = 1.6$  volts, which is in reasonable agreement with the observed value. In this calculation we used

the same numerical values as for  $\text{Cl}_2$  for the two first members of equation (12). This is equivalent to saying that we disregarded the deformation of the  $\text{S}^-$  ion by the  $\text{K}^+$  ion, and the possible differences in the collision diameter  $r_1$  and the size of the atoms. As to the electron affinity of the S atom, in respect to the first electron, it is rather uncertain; van Arkel and de Boer (*loc. cit.*, p. 64) assume the electron affinity of the oxygen atom to be approximately equal to that for halogens, i.e., about 4 volts. We assumed the same value also for sulphur. This interpretation seems to be inconsistent with the results obtained with ozone gas. According to the equation

$$(\text{O}_3) = (\text{O}_2) + (\text{O}) - 1.04 \text{ volts}, \quad (30)$$

the calculation gives  $E_m = 3.7$  volts and the observed value was about 1.5 volts. The disagreement is greater than that which can be allowed for the uncertainty of the calculations.

It is also necessary to point out that the phenomenon with COS has two features which differentiate it sharply from the reactions which result in the adsorption of halogens. In the first place the yield at pressures  $\leq 10^{-5}$  mm. was immeasurably small, i.e., the current was less than  $10^{-14}$  amp. In the second place the relation between negative emission and gas pressure is quite different. There is no sharp rise to an acute maximum at a pressure below  $10^{-4}$  mm. followed by a falling off at higher pressures but the emission rises at a steadily diminishing rate up to a maximum value at about  $10^{-3}$  mm. after which it remains constant as the pressure is increased further. It will be noticed from fig. 7 that the emission does not seem to start from  $p = 0$  but from about  $p = 10^{-4}$  mm. It is improbable that this is of much significance. It is more likely to be due to the fact that the absolute values of the pressure are not known accurately, owing to the Pirani gauge having been calibrated only in  $\text{N}_2$ , coupled with large errors in the determination of the very small currents involved at these low pressures.

Our other experiment with  $\text{N}_2\text{O}$ , made in order to investigate the effect of the oxygen atom in the reaction with a similar simple mechanism of dissociation

$$(\text{N}_2\text{O}) = (\text{N}_2) + (\text{O}) - 1.7_8 \text{ volts}, \quad (31)$$

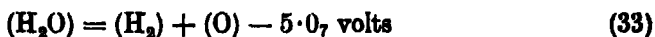
but with a much lower energy of dissociation, was unsuccessful. No electron emission was observed and there was no evidence that any chemical action took place. The glitter of the metal drops remained unchanged and the alloy surface in the lower bottle was absolutely clear. We had some reason, however, to expect such a result, namely, the very considerable inertia of the  $\text{N}_2\text{O}$  mole-

cules observed in some homogeneous changes. According to Hinshelwood and Burk\* 58.5 k. cal. are required to make the two gram molecules of  $N_2O$  enter into the gaseous bimolecular exothermic reaction

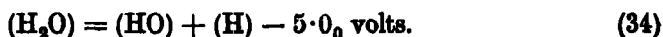


This suggests that the  $N_2O$  molecules have completely saturated valencies, and probably because of this there was no such adsorption of the  $N_2O$  molecules on the alkali metal as would result in the formation of a polar bond.

It is evident from the results of the present experiments that the emission, when the alloy  $NaK_3$  is acted on by  $H_2O$  gas in the normal state, is very small. It was at most  $25 \times 10^{-14}$  amp. even at pressures as high as about  $10^{-3}$  mm. The maximum currents with  $Cl_2$  were about  $10^5$  times and those with  $I_2$  about  $2 \times 10^3$  times this value. This result is consistent with the others when we consider the values of chemical energy resulting from the different alternative mechanisms of the reaction. Disregarding the possibility of the three atoms of the  $H_2O$  molecule simultaneously entering into the reaction with three metal atoms, there are two possible mechanisms of dissociation



and



As it is very unlikely that the  $(H_2)$  molecule will enter into the reaction with the alloy, we should expect a very small yield of the emission resulting from the adsorption of the O atoms because even in the reaction with ozone, which corresponds to a much lower diasociation energy, the yield was only about  $0.004 \times 10^{-4}$  amp./mm. The dissociation process (34) may either give a one-stage or a two-stage reaction. As the electron affinity of the pseudo-haloid (OH) is about the same as for halogens, and that of the (H) atom is only about 0.7 volt, the reaction will most probably go in two stages (*cf.* § 6). In such a case the first stage will give a negative value for  $E_m$  and only the second stage, which corresponds to the adsorption of the free (H), or (OH), atoms, will give a positive value for  $E_m$  but, as we know, the yield of the second stage of a two-stage reaction will be small.

It should be mentioned that one of us† found an emission of amount much more comparable with that given by  $Cl_2$  and  $COCl_2$  when the vapours, presumed to be  $H_2O$ , given off by  $CaCl_2 \cdot 6H_2O$  crystals, or by concentrated sulphuric

\* 'Proc. Roy. Soc.,' A, vol. 106, p. 284 (1924).

† Richardson, 'Phil. Trans.,' A, vol. 222, p. 1 (1921).

acid, reacted with a liquid alloy of sodium and potassium. To explain the results of those experiments it is necessary to assume either that the  $\text{H}_2\text{O}$  vapour given off from these reagents is in some kind of an abnormal state or that they were contaminated with some more active gas such as chlorine. As only a very small quantity would be required to cause the observed effects this is not improbable.

### § 9. *The Yield of Electrons.*

We can take as a measure of the yield of electrons in the reactions the limiting value of the ratio current  $i_0$  in amperes to pressure  $p$  in millimetres of mercury, as  $p$  tends to zero. The values of this quantity for the reactions investigated in the present paper are set out in the last column of Table VIII. The quantity  $D$  in the second column is the part of the total available energy which is lost, in the course of the reaction, to the molecular structure from

Table VIII.

Gas.	D in k. cal.	$E_0$ in volts, calculated.	$E_m$ in volts, calculated.	$E_m$ in volts, observed.	Yield $\times 10^4$ in amp./mm.
$\text{Cl}_2$	29	5.7 <sub>7</sub>	3.2 <sub>7</sub>	2.8 <sub>0</sub>	20 (25)
$\text{NOCl}$	38	5.4 <sub>2</sub>	2.9 <sub>2</sub>	2.6 <sub>0</sub>	7 (10)
$\text{COCl}_2$	42	5.2 <sub>2</sub>	2.7 <sub>2</sub>	2.5 <sub>2</sub>	2 (3)
$\text{HCl}$	(102)	(2.6 <sub>4</sub> )	(0.1 <sub>4</sub> )	—	$\nabla$ 0.0002
$\text{Br}_2$	45	4.5 <sub>0</sub>	2.0 <sub>0</sub>	1.7	2 (5)
$\text{I}_2$	35	4.3 <sub>2</sub>	1.8 <sub>2</sub>	1.5	0.3 (0.4)
$\text{COS}$	73	(4.1)	(1.1)	ca. 1.0	ca. 0.00003
$\text{H}_2\text{O}$	(115)	—	—	—	$\Delta\Delta$ 0.000003
$\text{N}_2\text{O}$	41	—	—	no reaction	$\Delta\Delta$ 0.000002

which the electron derives it. Thus in the one-stage reaction with  $\text{Cl}_2$  it is half the dissociation energy of  $\text{Cl}_2$ , in the first stage of the two-stage reaction with  $\text{Br}_2$  it is the whole of the dissociation of  $\text{Br}_2$ , in the one-stage reaction with  $\text{NOCl}$  it is the heat of the reaction  $(\text{NOCl}) = (\text{NO}) + (\text{Cl})$ , and so on. The values of  $E_0$  the chemical energy available in the reaction which can be given to an electron are got by the method of § 6, Part II, where the application to  $\text{COCl}_2$  is worked out. The calculated maximum energy  $E_m$  of the emitted electrons in the next column is got by subtracting 2.5 volts, the work function of  $\text{NaK}_2$ , from the corresponding value of  $E_0$  in the preceding column. The numbers under  $E_m$  observed are the measured values of the maximum energies of the main group of electrons which are obtained when the "tails" are

disregarded. This circumstance might easily explain why the observed values of  $E_m$  are a little too low, as compared with the calculated values of  $E_m$ , though the approximative character of the method employed in the calculations does not permit a very close comparison.

A very striking feature of Table VIII is the very large variation of the yield in the different reactions. Excluding  $N_2O$  in which there was no evidence that any chemical action was occurring the yield in the different reactions varies by a factor of nearly  $10^7$ . With  $H_2O$ , where it was clear that the chemical reaction was going on, we cannot say with certainty that there was any electron emission, i.e., that the yield was not zero.

It will also be seen that the order of the reactions according to electron yield is also nearly the same as the order for maximum electron energy and for chemical energy available. It is strictly the same if we keep the groups of reactions which have something homologous about them, for example, the group  $Cl_2$ ,  $NOCl$ , and  $COCl_2$  which all go in one stage to a considerable extent and which all result in the adsorption of the chlorine atoms. The change in the yield is evidently a very sensitive function of the change in the energy involved. These points are also clear when we compare the homologous two-stage reactions with  $Br_2$  and  $I_2$ . We can, in addition, state broadly that every one of these reactions which it is certain goes with great energy gives rise to a large yield of electrons, and that every one of them which it is certain goes with low energy gives rise to a small yield of electrons.

The values of the yields cannot be given with great accuracy, as the absolute values of the pressures for the different gases may be uncertain by a factor of 2. It should be emphasized that the gas pressures measured on the Pirani gauge in these investigations are not the true pressures of the gas; they are the pressures of ( $N_2$ ) which would have the same effect on the Pirani gauge as the actual gas had at the pressure employed in the experiment (see Part I, p. 31). Some values which we believe to be the highest possible values are given in brackets in the last column of Table VIII. The highest value of the yield,  $25 \times 10^{-4}$  amp./mm. for  $Cl_2$ , is equivalent to the emission of  $1.5 \times 10^{-4}$  electrons for each adsorbed atom of chlorine. Similar data can be worked out for the other reactions by combining this with the data in the last column of Table VIII and the molecular constants of the relevant gases involved.

We propose to return to the question of the relation between yield and energy later, when the results of more reactions have been described.

In concluding we should like to take this opportunity to acknowledge our indebtedness to the Department of Scientific and Industrial Research for a grant which has made it possible to carry out those investigations.

*Summary.*

Characteristic energy distribution curves, and other data are given for the electrons emitted when  $\text{NaK}_2$  is acted on by  $\text{Cl}_2$ ,  $\text{NOCl}$ ,  $\text{COCl}_2$ ,  $\text{Br}_2$ ,  $\text{I}_2$ , and  $\text{COS}$  at very low pressures. With  $\text{HCl}$  and  $\text{H}_2\text{O}$  the emission is small but probably real, with  $\text{N}_2\text{O}$  very small and probably zero. For  $\text{N}_2\text{O}$  there was no evidence of chemical action. Current and pressure curves are given for  $\text{Cl}_2$ ,  $\text{NOCl}$ , and  $\text{COS}$  and current and time curves for  $\text{Cl}_2$ .

The data are discussed, and analysed in the light of the theory of the electron emission mechanism developed in Part II and found to be consistent with it. In particular the measured maximum electron energies are found to be equal to the differences between the chemical energy available in the reaction and the work function of the alloy, provided the reactions with  $\text{Cl}_2$ ,  $\text{NOCl}$ , and  $\text{COCl}_2$  go to a great extent in one stage, and those with  $\text{Br}_2$  and  $\text{I}_2$  go mainly in two stages. With the reactions which involve oxygen and sulphur either ambiguities in the interpretation, or gaps in the data, make similar simple statements impossible.

The electron yields are given and shown to be high for energetic and low for weak reactions. For homologous reactions the electron yield is a sensitive function of the energy available.



## *Two-Dimensional Oscillations in Divergent Jets of Compressible Fluid.*

By S. G. HOOKER, A.R.C.S., Brasenose College, Oxford.

(Communicated by R. V. Southwell, F.R.S.—Received December 19, 1933.)

1. The problem of determining the possible modes of stationary oscillation for a compressible fluid, moving with a steady velocity which is not constant, usually presents great difficulties. One case which is to some extent amenable to analysis is that of uniform radial flow in two dimensions, where the undisturbed paths of the fluid particles are straight lines radiating from a common point or source. The term "source" is here used somewhat loosely, for in the solution which will be given it is found that the fluid density attains unreal values inside a certain circle having its centre at this common point.

It is well known that in radial flow two systems of velocity are possible to a compressible fluid—namely, either

- (i) zero velocity at  $r = \infty$ , and an increasing but limited velocity as the radial distance from the source decreases (i.e., a modified "perfect fluid" motion); or
- (ii) the maximum possible velocity at infinity (corresponding with zero pressure and density), and a *decreasing* speed—also limited—as  $r$  decreases. This type of flow is peculiar to compressible fluids.

In both systems the limiting value of the radius, for which the fluid has a real density, is that at which the velocity of sound is attained. Since the geometry of the stream lines implies that the stream tubes are diminishing in cross-section as they approach the source, this feature exemplifies Osborne Reynolds' conclusion that the section of a stream tube is a minimum where the speed of sound occurs.

2. In the undisturbed state, the contours of constant velocity are concentric circles centred at the source, and the stream lines are the curves  $\theta = \text{const.}$  Let  $\psi$  be the stream function defined in relation to mass-flow, so that

$$-\frac{1}{\rho} \frac{\partial \psi}{r \partial \theta} \quad \text{and} \quad + \frac{1}{\rho} \frac{\partial \psi}{\partial r}$$

are the radial and circumferential components of velocity at  $(r, \theta)$ . The

equation of continuity requires that the same mass of fluid shall cross each concentric circle, and thus may be written in the form

$$2\pi r \frac{\partial \psi}{r \partial \theta} = \text{const.}$$

or

$$\psi = A\theta.$$

The constant  $A$  is fixed by the units of velocity and length. If we take the unit of velocity to be the local speed of sound, and assume that this speed occurs at  $r = 1$ , then

$$\psi = \theta$$

represents the stream function for the undisturbed flow. Let  $u_0$  and  $\rho_0$  represent the values of the radial velocity and density at distance  $r$ ; then the equation of continuity can be written in the form

$$2\pi r \rho_0 u_0 = \text{const.}$$

and if we take units such that

$$\rho_0 = u_0 = 1, \quad \text{when} \quad r = 1,$$

this becomes

$$r \rho_0 u_0 = 1 \tag{1}$$

3. Now the density  $\rho$  and the resultant velocity  $q$  of a compressible liquid satisfying the adiabatic law

$$p = A\rho^\gamma$$

are related by the formula

$$\rho = \{1 - \frac{1}{2}(\gamma - 1)(q^2 - 1)\}^{1/\gamma-1}, \tag{2}$$

where  $\rho$  and  $q$  are measured as fractions of the density and velocity corresponding with the speed of sound. A graph of this function is shown in fig. 1, in which  $\gamma$  has been given its value for air (1.408).

In purely radial flow,  $q = u_0$ , and combining (1) and (2) we can eliminate  $\rho_0$  and obtain a relation between  $r$  and  $u_0$ ; from this the values given in Table I can be calculated.

Table I.

$u_0$	0	0.2	0.4	0.6	0.8	1.0	1.2	1.5	2.0	2.45
$r$	$\infty$	3.22	1.70	1.23	1.05	1.0	1.05	1.37	4.95	$\infty$

This curve of  $u_0$  against  $r$  is shown in fig. 2; it demonstrates the existence of two possible systems of velocity outside the circle of minimum radius  $r = 1$ . For values of  $r < 1$  the velocity and density are not real.

4. Continuing the discussion of disturbed radial flow, we shall hereafter employ the stream function  $\psi$ , thereby automatically ensuring the satisfaction

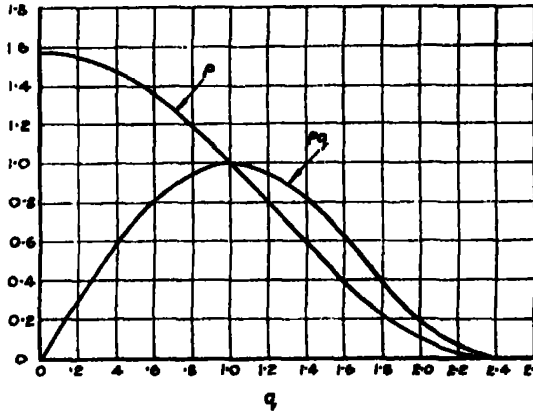


FIG. 1.—Curves relating density with velocity.

of the equation of continuity. Let us now assume that in a disturbed but steady state of flow

$$\psi = \theta + \Lambda f(r) g(\theta), \quad (3)$$

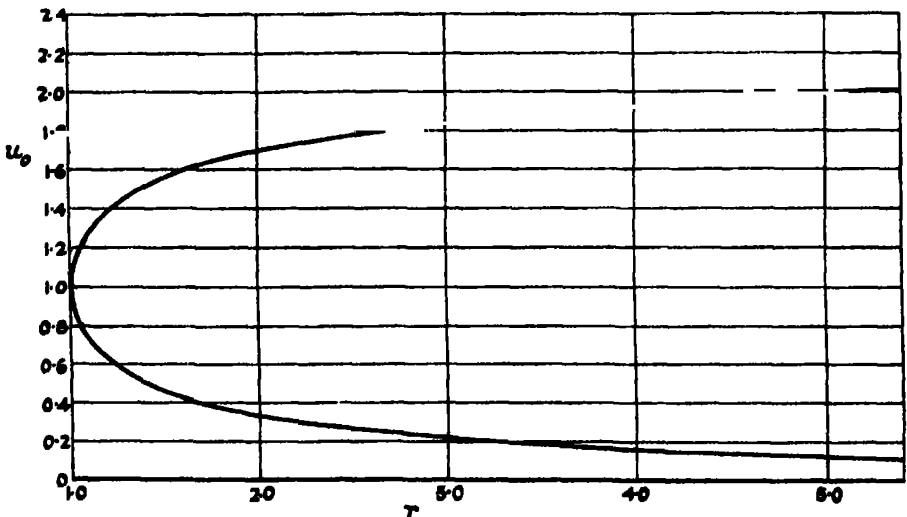


FIG. 2.—The undisturbed radial velocity.

where  $\Lambda$  is a small constant whose square can be neglected, and  $f(r)$  and  $g(\theta)$  are functions of  $r$  and  $\theta$  respectively, to be determined from the condition that the motion must remain irrotational.

If  $u$  and  $v$  are the radial and circumferential components of velocity, this condition is

$$\frac{\partial}{\partial r}(vr) - \frac{\partial u}{\partial \theta} = 0. \quad (4)$$

Now from (3), we have

$$\begin{aligned} \rho v &= \frac{\partial \psi}{\partial r} = A f'(r) g(\theta), \\ -\rho u &= \frac{\partial \psi}{r \partial \theta} = \frac{1}{r} + \frac{A f(r)}{r} g'(\theta). \end{aligned}$$

We substitute these values for  $u$  and  $v$  in (4), and thus obtain

$$\frac{\partial}{\partial r} \left\{ \frac{1}{\rho} A r f'(r) g(\theta) \right\} + \frac{\partial}{\partial \theta} \left\{ \frac{1}{\rho} \left( \frac{1}{r} + \frac{A f(r)}{r} g'(\theta) \right) \right\} = 0. \quad (5)$$

If  $A^2$  is negligibly small,

$$\rho q = |\rho u| = \frac{1}{r} + \frac{A f(r)}{r} g'(\theta), \quad (6)$$

and hence

$$\frac{1}{\rho} = \frac{qr}{1 + A f(r) g'(\theta)}.$$

Therefore

$$\frac{A r f'(r) g(\theta)}{\rho} = q r^2 A f'(r) g(\theta), \text{ to the first order in } A,$$

and

$$\frac{1}{\rho r} \{1 + A f(r) g'(\theta)\} = q.$$

Accordingly (5) can be written in the form

$$\frac{\partial}{\partial r} \{q r^2 A f'(r) g(\theta)\} + \frac{\partial q}{\partial \theta} = 0. \quad (7)$$

Now in the disturbed state

$$q = u_0 + \delta q,$$

where  $u_0$  is the undisturbed radial velocity (a function of  $r$  only), and  $\delta q$  is a function of  $r$  and  $\theta$ . Hence

$$\rho q = \rho_0 u_0 + \frac{\partial}{\partial u_0} (\rho_0 u_0) \delta q,$$

$\rho_0$  being the fluid density in the undisturbed flow. But, by virtue of continuity in the undisturbed state,

$$\rho_0 u_0 = \frac{1}{r}.$$

Therefore

$$\begin{aligned}\rho q &= \frac{1}{r} + \frac{\partial}{\partial u_0} (\rho_0 u_0) \delta q, \\ &= \frac{1}{r} + \frac{A f(r)}{r} g'(\theta), \quad \text{by (6),}\end{aligned}$$

and we have

$$\delta q = \frac{A f(r) g'(\theta)}{r \frac{\partial}{\partial u_0} (\rho_0 u_0)}.$$

Since  $\delta q$  is the only part of  $q$  involving  $\theta$ , equation (7) becomes

$$\frac{\partial}{\partial r} \{u_0 r^2 f'(r) g(\theta)\} + \frac{f(r) g''(\theta)}{r \frac{\partial}{\partial u_0} (\rho_0 u_0)} = 0,$$

which can be written in the form

$$\frac{\left[ \frac{d}{dr} \{u_0 r^2 f'(r)\} \right] r \frac{\partial}{\partial u_0} (\rho_0 u_0)}{f(r)} = - \frac{g''(\theta)}{g(\theta)},$$

or

$$(\text{Function of } r) = (\text{Function of } \theta),$$

which is possible only if both sides are constant. Hence  $g(\theta)$  must be a periodic function, and we can write

$$\frac{g''(\theta)}{g(\theta)} = -m^2,$$

where  $m$  is a real number, so that

$$g(\theta) = \begin{cases} \cos m\theta, \\ \sin m\theta. \end{cases} \quad (8)$$

Also we have

$$\frac{d}{dr} \{u_0 r^2 f'(r)\} - \frac{m^2 f(r)}{r \frac{\partial}{\partial u_0} (\rho_0 u_0)} = 0,$$

and, since  $\rho_0 u_0 r = 1$ , this equation can be written in the form

$$\frac{r}{\rho_0} \frac{d}{dr} \left\{ \frac{r}{\rho_0} \frac{df}{dr} \right\} - \frac{m^2 f(r)}{\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0)} = 0. \quad (9)$$

If we now take a new independent variable  $n$ , such that

$$\frac{\rho_0 dr}{r} = dn. \quad (10)$$

equation (9) takes the form

$$\frac{d^2 f}{dn^2} - \frac{m^2 f}{\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0)} = 0, \quad (11)$$

in which  $\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0)$  must now be expressed as a function of  $n$ .

5. Equation (11) is of a type which also governs the modes of deflection of a strut of variable cross-section under a compressive end loading  $P$ ,—viz.,

$$B \frac{d^2 y}{dx^2} + Py = 0,$$

where  $B$  ( $= EI$ ) is the flexural rigidity. The two equations are identical if

$$P/B = -m^2 / \rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0),$$

and the end conditions representing a "simply supported" rod,—namely,  $y = 0$  at both ends—are required by the solution of (11), since it is necessary that  $f = 0$  at both ends of the range of  $n$ .

Thus we could obtain the solution of (11) by considering the stability of a rod under an end loading  $P$  given by

$$P = -m^2 EI / \rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0).$$

This end loading will be *tensile* if

$$\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0) > 0,$$

i.e., if  $\frac{\partial}{\partial u_0} (\rho_0 u_0) > 0$ , since  $\rho$  is essentially positive.

The rod will obviously be stable under tension; and correspondingly there will be no oscillations of the fluid if

$$\frac{\partial}{\partial u_0} (\rho_0 u_0) > 0.$$

This condition corresponds with  $u_0 < 1$ , and as a consequence we need look for oscillatory solutions only when the velocities are greater than that of sound.

6. To determine the range of  $n$  corresponding with the full range of  $r$ ,—i.e.

$$1 \leq r < \infty,$$

we observe that the range of  $u_0$  in which we are now interested is such that

$$1 \leq u_0 \leq \sqrt{\frac{\gamma+1}{\gamma-1}},$$

—the latter limit being the velocity at which the fluid density and pressure become evanescent, and so corresponding with ( $r = \infty$ ).

Writing (10) in its alternative form

$$n = \int_1^r \frac{dr}{u_0 r^2}, \quad (12)$$

we see that if  $N$  is the total range of  $n$ ,

$$N < \int_1^\infty \frac{dr}{r^2} \quad (\text{giving } u_0 \text{ its lower limit over the whole range}),$$

and

$$N > \sqrt{\frac{\gamma-1}{\gamma+1}} \int_1^\infty \frac{dr}{r^2} \quad (\text{giving } u_0 \text{ its upper limit over the whole range}).$$

The value of the integral being unity, it follows that

$$\sqrt{\frac{\gamma-1}{\gamma+1}} < N < 1.$$

For air, for which  $\gamma = 1.408$ , this becomes

$$0.408 < N < 1.$$

Table II.

$r$	1.0	1.1	1.2	1.4	1.6	1.8	2.0	2.4
$n = \int_1^r \frac{dr}{u_0 r^2}$	0	0.076	0.132	0.212	0.268	0.308	0.340	0.388
$r$	2.8	3.2	3.6	4.0	5.0	8.8	19.75	87.5
$n = \int_1^r \frac{dr}{u_0 r^2}$	0.420	0.447	0.469	0.487	0.517	0.573	0.613	0.654

To determine the exact range of  $n$  we proceed as follows. In fig. 2,  $u_0$  has been given as a function of  $r$ . From this we can plot  $1/u_0 r^2$ , and hence by graphical methods determine the value of the integral (12). When this is done the values given in Table II are obtained.

The value to which  $n$  tends can be determined more definitely by plotting  $u_0$  against  $n$ , using the values in Table II.

Now

$$\frac{du_0}{dn} = \frac{du_0}{dr} \cdot \frac{dr}{dn},$$

and we observe that  $du_0/dr \rightarrow \infty$ , when  $n \rightarrow 0$ , while as  $n$  attains its limiting finite value  $dr/dn \rightarrow \infty$ . Hence the curve relating  $u_0$  and  $n$  has the form given in fig. 3: the limiting value of  $u_0$  being 2.45, the estimated range of  $n$  from this figure is

$$0 \leq n \leq 0.68.$$

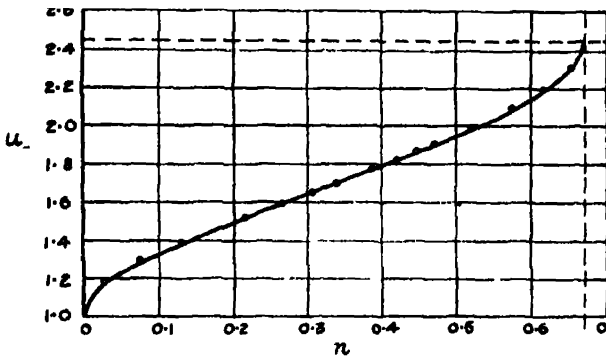


FIG. 3.

7. It does not appear possible to express the function

$$\frac{1}{\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0)},$$

occurring in (11), as an analytic function of  $n$ . A graphical representation can, however, be obtained by simple, if somewhat lengthy, methods.

In fig. 1 the graph of  $\rho$  against  $q$  was given: on the same diagram we can plot the curve of  $\rho q$  against  $q$ . By graphical differentiation of this curve we obtain the graph of  $\frac{\partial}{\partial q} (\rho q)$  against  $q$ , fig. 4. Since it is difficult to obtain accuracy with graphical differentiation, the derived values were tested by reintegration.

We are now in the position to plot  $\rho \frac{\partial}{\partial q} (\rho q)$  against  $q$ , and with the assistance of fig. 3 (determining  $q$  in terms of  $n$ ) we can obtain the graph of the function  $\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0)$  against  $n$ . This is shown in fig. 5.



At the central position ( $n = 0.34$ ) the value of the function is  $-0.29$ . Taking a new independent variable  $N = n/0.68$ , we can plot finally the function

$$\frac{I}{I_0} \equiv - \frac{0.29}{\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0)} \quad \text{against } N \left( \equiv \frac{n}{0.68} \right).$$

This is given in fig. 6.

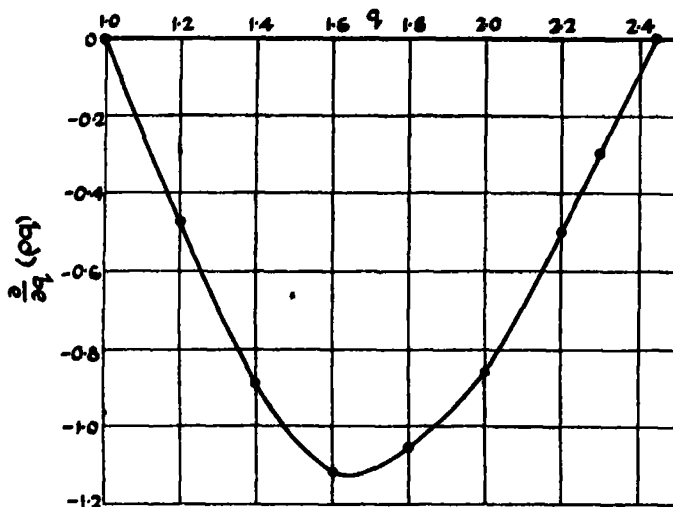


FIG. 4.—The derived curve.

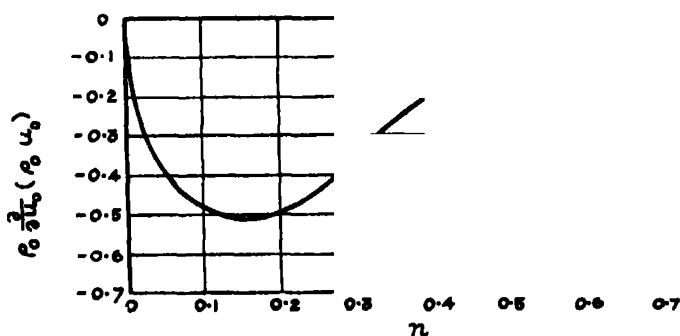


FIG. 5.—Graph of the function  $\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0)$  plotted against  $n$ .

8. Reverting now to the differential equation (11), we see that

$$\frac{d^2 f}{dn^2} - m^2 I f = 0,$$

or (putting  $N = n/0.68$ )

$$\frac{1}{(0.68)^2} \cdot \frac{d^2 f}{dN^2} + m^2 \frac{1}{I_0} f \times \frac{1}{0.29} = 0,$$

or

$$\frac{d^2 f}{dN^2} + \frac{m^2}{0.625} \frac{1}{I_0} f = 0. \quad (13)$$

The integrations of (13) were performed graphically for the following values of  $m$  :—

$$m = 0, 2, 4, 10.$$

$1/I_0$

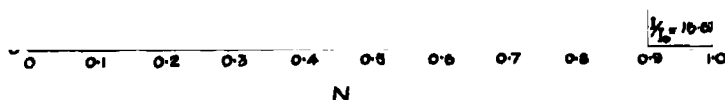


FIG. 6.

For  $m = 0$ , (13) reduces to

$$\frac{d^2 f}{dN^2} = 0,$$

whence  $f = N$  is the admissible solution. For  $m = 2, 4$ , or  $10$ , the method of procedure was as follows :

The second-order differential equation (13) is split up into two simultaneous first-order equations :—

$$\frac{dg}{dN} = - \frac{m^2}{0.625} \frac{1}{I_0} f \quad (a)$$

and

$$\frac{df}{dN} = g. \quad (b)$$

Since the absolute magnitude of  $f$  is not defined by equation (13), we may take  $f$  to have any convenient slope at the origin, and there is no loss of generality if we make it unity. Taking the tangent as a sufficient representation of the curve over a small interval  $N = \delta$ , we obtain  $f = \delta$  when  $N = \delta$ . Substituting this value for  $f$  in (a), and then for a given value of  $m \neq 0$ ,  $dg/dN$  can be calculated. Plotting this, we can obtain the value of  $g = \int_0^\delta \frac{dg}{dN} dN$ ,

and so from (b) determine the new slope of  $f$  at  $N = \delta$ . Continuing the curve with this calculated slope for a further distance  $\delta$ , we can repeat the process, and by a step-by-step method traverse the whole curve.

In practice the method works very easily and quickly. The integrations to obtain  $g$  can be performed with sufficient accuracy by the mean ordinate method, so that actually there is no need to plot  $dg/dN$ . There is the added advantage that at any moment the steps can be shortened or lengthened as the prevailing circumstances may demand.

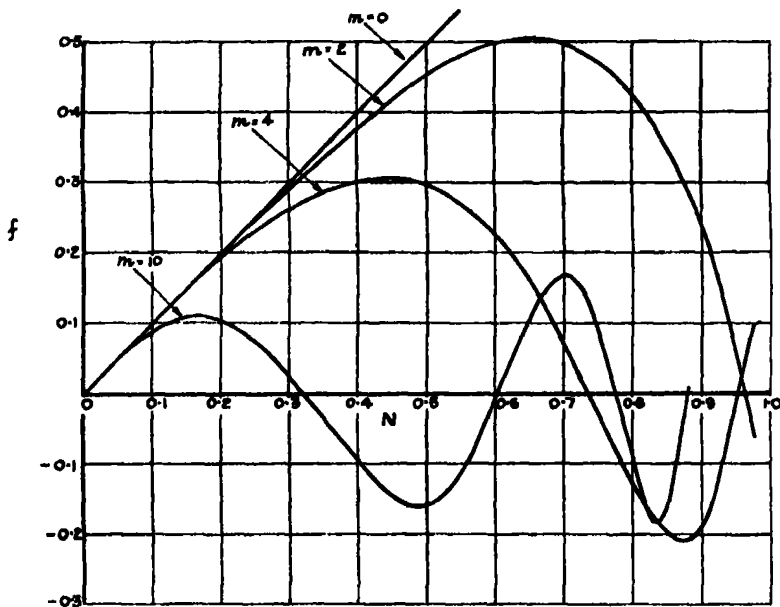


FIG. 7.

9. A summary of the calculations for the case  $m = 10$  is given in the Appendix, p. 70, and in fig. 7 the final curves for the function  $f$  are shown. Using the preceding method, the solution can be carried (by taking small enough steps) as near to  $N = 1$  as may be desired; but (since  $N = 1$  corresponds with  $r = \infty$ ) there is little interest in so doing.

On the other hand, it can be shown that the solution of equation (13) has only a finite number of roots in the range  $0 \leq N < 1$ .\* For if there were an infinite number of zeros, these zeros would have at least one limit point  $N = n_0$  at which not only

$$f(n_0) = 0,$$

but

$$f'(n_0) = 0.$$

But it is evident from (13) that where

$$f(n_0) = 0,$$

then

$$f''(n_0) = 0.$$

Hence, by successive differentiations of (13), we can show that at a limit point all the differential coefficients vanish, and in consequence the function  $f$  must be identically zero over the whole range.

10. Finally, the solutions can be plotted as functions of  $r$ ; as such, they are shown in fig. 8.

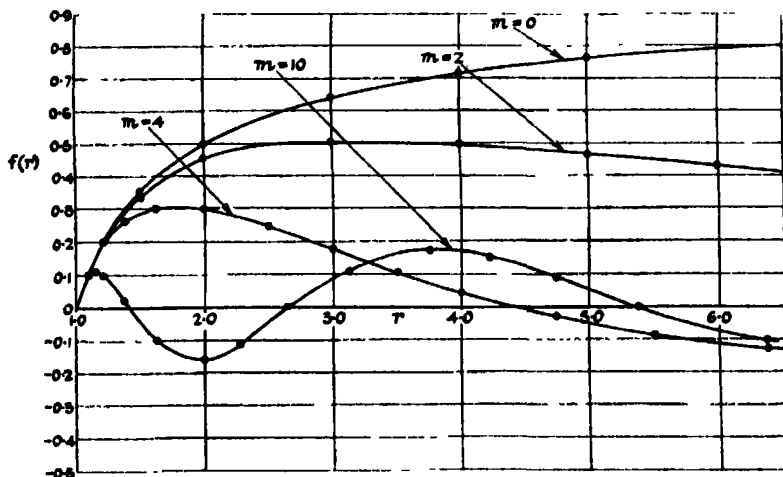


FIG. 8.—The whole angle of the jet is  $\pi/m$ .

Taking

$$\psi = \theta + A f(r) \cos m\theta,$$

as the disturbed form of  $\psi$ , we find that

$$\rho v = A f'(r) \cos m\theta,$$

$$-\rho u = \frac{1}{r} - \frac{mA f(r)}{r} \sin m\theta.$$

\* See Ince, "Ordinary Differential Equations," p. 223.

It follows that  $v = 0$  when either

$$\text{or } \left. \begin{array}{ll} \text{(i)} & m\theta = \pm \pi/2 \\ \text{(ii)} & f'(r) = 0. \end{array} \right\} \quad (14)$$

From condition (i) we see that the circumferential velocity  $v$  is zero at all points of the lines  $\theta = \pm \pi/2m$ , and hence these lines of flow are not distorted by the disturbance. They can, in fact, be regarded as the solid boundaries of a nozzle, and then we can consider the disturbance defined by  $\psi$  as being confined to a uniformly divergent jet of semi-angle  $\pi/2m$ , and as consisting of oscillations which are antisymmetrical with respect to the axis of the jet. The case of symmetrical oscillations is treated later in § 15.

In the special case when  $m = 0$ ,

$$\psi = 0 + Af(r).$$

With this form for  $\psi$ , both  $u$  and  $v$  are independent of  $\theta$  and are functions of  $r$  only. There are therefore no undisturbed stream lines, and the distortion of each is the same. The corresponding motion is thus of a spiral type.

A general idea of a typical stream line in all cases can be got from fig. 8, in which the axis  $f(r) = 0$  is to be regarded as the undisturbed line. For at the points where  $f'(r) = 0$  the curve of  $f(r)$  is moving parallel to the axis; and at these positions, from (14), we notice that  $v = 0$ , so that stream lines are moving sensibly parallel to their undisturbed positions.

It has already been shown that in all cases there are only a finite number of zeros of the function  $f$  in the range  $0 < N < 1$ . Hence a stream line only makes a finite number of oscillations in going from  $r = 1$  to  $r = \infty$ , with the result that the "wave-length" must tend to become infinite as the speed of flow tends to its limiting maximum value. This is in accordance with Prandtl's\* theory of the uniform jet, in which it can be shown that the wave-length of a possible stationary oscillation tends to become infinitely long as the speed of the stream tends to the limit at which the pressure and density of the fluid become evanescent.

The maximum change in shape of the stream lines always occurs near  $r = 1$ , i.e., near the point at which the speed of sound is attained. Figs. 7 and 8 show that as  $m$  increases (i.e., as the angle of the nozzle becomes smaller) the oscillations become more frequent and tend to become concentrated near ( $r = 1$ ). It appears that in nozzles of small angle the fluid may be highly disturbed in the region where the speed of sound occurs.

\* 'Phys. Z.', vol. 5, p. 599 (1904); vol. 8, p. 23 (1907).

11. In this connection we may examine the form of the solution of (13) when  $m$  becomes very large and  $n$  is small. Near the velocity of sound, it is evident from fig. 1 that when

$$q = 1 + \Delta q,$$

then

$$\rho = 1 - \Delta q.$$

Hence  $\rho q = 1 - (\Delta q)^2$ ; and it follows from equation (1) that for radial flow near the speed of sound

$$u_0 = 1 + \Delta q = 1 + \sqrt{1 - \frac{1}{r}},$$

$$\rho_0 = 1 - \Delta q = 1 - \sqrt{1 - \frac{1}{r}}.$$

Putting  $t = \sqrt{1 - \frac{1}{r}}$ , we see that

$$u_0 = 1 + t \quad \text{and} \quad \rho_0 = 1 - t.$$

Hence

$$\begin{aligned} dn &= \frac{\rho_0 dr}{r} = \frac{2t dt}{(1 - t^2)^2} (1 - t^2) (1 - t) \\ &= \frac{2t dt}{1 + t}, \end{aligned}$$

while

$$\rho_0 \frac{\partial}{\partial u_0} (\rho_0 u_0) = -2t(1 - t).$$

In these circumstances equation (13) becomes

$$(1 - t^2) \frac{d}{dt} \left( \frac{1 + t}{2t} \frac{df}{dt} \right) + m^2 f = 0. \quad (15)$$

In (15) write

$$t' = mt,$$

and the equation becomes

$$\left( 1 - \frac{t'^2}{m^2} \right) m \frac{d}{dt'} \left( \frac{m}{2t'} + \frac{1}{2} \right) m \frac{df}{dt'} + m^2 f = 0.$$

Neglecting  $t'^2/m^2$  and  $\frac{1}{2}$ , we have

$$\frac{d}{dt'} \left( \frac{1}{2t'} \frac{df}{dt'} \right) + \frac{f}{m} = 0.$$

Differentiate this equation with respect to  $t'$  and write

$$\chi = \frac{1}{2t'} \frac{df}{dt'} \left( \equiv \frac{1}{m^2} \frac{1}{2t'} \frac{df}{dt'} \right).$$

Then we have

$$\frac{d^2\chi}{dt'^2} + \frac{2t'}{m}\chi = 0. \quad (16)$$

Writing  $t' = \beta\eta$ , equation (16) can be thrown into the form

$$\frac{d^2\chi}{d\eta^2} + \frac{2}{\beta}\chi\eta = 0, \quad (17)$$

provided that  $\beta$  satisfies the relation  $2\beta^3/m = \frac{9}{4}$ .

12. The complete solution of (17) is known to be

$$\chi = A\eta^{1/3}J_{1/3}(\eta^{3/2}) + B\eta^{1/3}J_{-1/3}(\eta^{3/2}),$$

and the conditions of the problem are satisfied by taking

$$\chi = B\eta^{1/3}J_{-1/3}(\eta^{3/2}).$$

In this expression, write

$$x = \eta^{3/2},$$

so that

$$\chi = Bx^{1/3}J_{-1/3}(x). \quad (18)$$

But

$$\chi = \frac{1}{2t'} \frac{df}{dt'} = \frac{1}{2\beta\eta} \frac{df}{\beta d\eta} = \frac{1}{2\beta^2\eta} \frac{df}{d\eta},$$

and expressing  $\frac{1}{2\beta^2\eta} \frac{df}{d\eta}$  in terms of the new independent variable  $x$ , we have

$$\frac{1}{2\beta^2\eta} \frac{df}{d\eta} = \frac{3}{4\beta^3} \frac{1}{x^{1/3}} \frac{df}{dx}.$$

Equating the latter expression for  $\chi$  to the right-hand side of (18), we obtain

$$\frac{df}{dx} = B' x^{2/3} J_{-1/3}(x),$$

which, by a well-known formula relating Bessel functions, can be written

$$\frac{df}{dx} = B' \frac{d}{dx} \{x^{2/3} J_{2/3}(x)\}.$$

Integrating this equation, the expression for  $f$  which vanishes when  $x = 0$  is

$$f = B' x^{2/3} J_{2/3}(x),$$

where  $x = \eta^{3/2}$ ,  $t' = \beta\eta$ ,  $t' = m\epsilon$ , and  $2\beta^3/m = \frac{9}{4}$ .

Making these changes to obtain  $f$  in terms of the original variable  $t$ , we obtain the final expression

$$f = B'' \epsilon J_{2/3} \left( \frac{2\sqrt{2}}{3} m\epsilon^{3/2} \right). \quad (19)$$

The expansion of  $J_{2/3}$  being

$$J_{2/3} \left( \frac{2\sqrt{2}}{3} m t^{2/3} \right) = \frac{\left( \frac{2\sqrt{2}}{3} m \right)^{3/2} t}{2^{2/3} \Pi \left( \frac{2}{3} \right)} \left\{ 1 - \frac{\frac{8}{3} m^2 t^2}{2 \left( \frac{4}{3} + 2 \right)} + \dots \right\},$$

by making

$$B'' = \frac{\Pi \left( \frac{2}{3} \right)}{\left( \frac{\sqrt{2}}{3} m \right)^{3/2}},$$

the expression for  $f$  as a power series becomes

$$f = t^2 \left\{ 1 - \frac{\frac{8}{3} m^2 t^2}{2 \left( \frac{4}{3} + 2 \right)} + \dots \right\},$$

and therefore vanishes to order  $t^2$  as required.

If the roots of  $J_{2/3}(x)$  are  $n_1, n_2, n_3, \dots$ , etc., then from (19)  $f$  vanishes when

$$\frac{2\sqrt{2}}{3} m t^{2/3} = n_s, \quad (s = 1, 2, 3, \dots, \text{etc.}).$$

By making  $m$  sufficiently large, we can make as many of these values of  $t$  as we please approach as close as desired to  $t = 0$ . Thus for large values of  $m$  there will be a large number of oscillations concentrated near  $t = 0$ , or near where the fluid has the speed of sound.

Furthermore the approximation employed for the relation between density and velocity—viz. :

$$\rho = 1 - \Delta q,$$

when

$$q = 1 + \Delta q$$

is independent of the value of  $\gamma$ , and holds equally well for air or steam. An examination of fig. 1 will show that it is sufficiently accurate for the range

$$1.0 \leq q \leq 1.6,$$

and thus the analytical solution given by (19) is suitable for the determination of the oscillations in short nozzles of small angle.

13. While no direct experimental evidence appears to be available to check the accuracy of the foregoing theory, it is borne out in a qualitative manner by the photographs given by Busemann.\* Three photographs are there shown of a jet escaping into a region of over, equal, and under back-pressures; oscillations are clearly defined at all three pressures, not only in the free jet, but also in the interior of the nozzle itself.

\* 'Handbuch der Experimental Physik,' vol. 4, 1 Teil, p. 429.



The photographs were taken by the Schlieren method, which depends for its success on variations of density in the fluid. From the photographs it is evident that the oscillations do not extend past the throat of the nozzle, but are confined to the region in which the velocity of sound is exceeded. Further, the "wave-length" increases rapidly in the region where the nozzle diverges.

14. Referring to the theoretical work, we see that the density will be affected only by  $u$  and not by  $v$ . As a consequence the maximum effect observable by the Schlieren method will occur where the disturbed addition to  $u$  is a maximum. This is where  $f(r)$  is a maximum,—a condition which always occurs at a value of  $r$  greater than 1 and at a speed in excess of that of sound. This again is borne out by the photographs already mentioned.

It is also of considerable interest to see that oscillations actually occur in the interior of the nozzle itself. Stodola\* considered that when vibrations do occur in the interior of the nozzle the fluid has detached itself from the wall. The preceding theoretical discussion shows that this condition is not necessary—at any rate for a nozzle of uniform divergence.

15. Finally we may consider the pressure variation along the axis of the nozzle. For oscillations which are antisymmetrical with respect to the central plane of the nozzle, the appropriate expression for  $\psi$  is

$$\psi = \theta + Af(r) \cos m\theta,$$

and we have

$$\rho v = Af'(r) \cos m\theta,$$

and

$$\rho u = \frac{1}{r} [1 + Af(r) \{-m \sin m\theta\}].$$

The axis of the nozzle is the line  $\theta = 0$ , along which

$$\rho u = 1/r$$

and

$$\rho v = Af'(r).$$

But to the first order in  $A$  the density (and therefore the pressure) is independent of  $v$ ; so to determine the pressure we need only obtain  $u$  from the relation

$$\rho u = 1/r.$$

But this is the relation holding in the undisturbed flow. Accordingly the pressure along the axis will not be sensibly affected by the presence of the oscillations.

\* "Steam Turbines," p. 85.

*It appears useless, therefore, to attempt to discover small oscillations of this type in a nozzle by measuring the pressure variation along the axis: it is at the wall, where the fluid is constrained to move in a straight line, that the maximum variation in pressure will occur. But once the jet has left the nozzle, it is the axis which becomes undeflected, while the maximum oscillation occurs at the outer edges; therefore a pressure investigation along the axis should now demonstrate the presence of the oscillations.*

In experiment the oscillations may be expected to exhibit symmetry with respect to the central plane of the jet. This case can be obtained from the preceding analysis by merely doubling the angle of the nozzle. The pressure variation along the axis, due to the oscillations, will evidently now be appreciable and equal to that obtaining at the walls.

16. In his experimental investigations on the flow from diverging nozzles, Stodola used a pressure-exploring tube mounted along the nozzle axis. The foregoing may serve to explain why he never found (except under very exceptional circumstances) oscillations occurring in the interior of the nozzle. Inevitably there was a total absence of any pressure vibration in the regular expansion line, but once the jet became free vibrations were immediately made manifest.

In concluding this paper, which has been written at the Engineering Laboratory, Oxford, I should like to express my gratitude to Professor R. V. Southwell, for his interest in and criticism of the work. I am also indebted to the Department of Scientific and Industrial Research for a Senior Research Award, which has made this investigation possible

### *Summary.*

In this paper the work of Prandtl and Rayleigh on stationary vibrations in uniform jets of elastic fluid is extended by a different method to oscillations in expanding jets.

The stream function  $\psi$  for the flow in a straight divergent nozzle being given by

$$\psi = 0,$$

it is assumed that in the disturbed condition

$$\psi = \theta + Af(r)g(\theta).$$

It is then shown that the function  $g(\theta)$  is of the form  $e^{im\theta}$ , and that we may regard the disturbance as being confined to a straight divergent nozzle of angle  $\pi/m$  radians.

The differential equation governing the form of  $f(r)$  is of the second order. The author was not able to solve it completely by analytical methods; but solutions were obtained by graphical methods for four selected nozzles of angle

$$360^\circ, 90^\circ, 45^\circ, \text{ and } 18^\circ.$$

For smaller angles the approximate equations were developed, and found to be soluble in terms of Bessel functions of fractional order.

Points of interest arising are :—

(a) As the angle of a nozzle becomes small the oscillations tend to be concentrated at the throat, near where the acoustic velocity occurs.

(b) It is shown that the pressure fall along the axis of the nozzle is unaffected by the existence of antisymmetrical stationary vibrations. The "search tube" method of demonstrating the presence of oscillations fails under these conditions; and the conclusion of Stodola, that oscillations occur only in the interior of nozzles when the flow has broken away from the walls, is therefore not necessarily true.

(c) Photographs included in Busemann's article in the "Handbuch der Physik" clearly show vibrations in the interior of nozzles, and thus support the theory in a qualitative manner.

#### APPENDIX.

Summary of the solution of  $\frac{d^2f}{dN^2} + \frac{m^2}{0.625} \frac{1}{I_0} f = 0$  for  $m = 10$ .

The second-order equation is split up into two simultaneous first-order equations, viz. : —

$$\frac{df}{dN} = g$$

$$\frac{dg}{dN} = -\frac{m^2}{0.625} \frac{1}{I_0} f.$$

A start is made by taking  $df/dN = 1.00$  when  $N = 0$ ; and the values of  $I/I_0$  are taken from fig. 6.

Step	1	2	3	4	5	6	7	8	9
N	0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
f	0	0.05	0.091	0.11	0.105	0.075	0.025	-0.037	-0.097
dg/dN	0	-6.97	-9.76	-10.55	-9.68	-6.6	-2.28	3.67	10.85
$g \equiv df/dN$	1.00	0.825	0.406	-0.095	-0.596	-1.003	-1.225	-1.190	-0.827

Step	10	11	12	13	14	15	16	17	18
N	0.45	0.475	0.50	0.525	0.55	0.575	0.625	0.65	0.675
$f$	-0.145	-0.16	0.16	-0.145	-0.11	-0.065	+0.05	0.105	0.15
$dg/dN$	18.6	23	25.6	25.5	21.1	13.7	-13.2	-31.9	-50.5
$g \equiv df/dN$	-0.509	+0.011	0.619	1.258	1.841	2.276	2.276	1.712	0.682

Step	19	20	21	22	23	24	25
N	0.70	0.725	0.75	0.80	0.825	0.84	0.85
$f$	0.167	0.15	0.092	-0.10	-0.175	-0.18	-0.16
$dg/dN$	-64.2	-67.2	-47.1	72	160	197	197
$g \equiv df/dN$	-0.752	-2.392	-3.822	-3.190	0.299	2.38	4.35

### *An Analysis of the Dipole Moments of some Aromatic Compounds.*

By GEORGE MACDONALD BENNETT and SAMUEL GLASSTONE, University of Sheffield.

(Communicated by C. K. Ingold, F.R.S.—Received December 20, 1933,  
Revised February 15, 1934.)

The dipole moments of organic substances represent the resultants of a number of distinct moments within the molecule from which they must, as was pointed out by Sir J. J. Thomson,\* be obtainable by vector summation. This view was extended and further examined by Eucken and Meyer in 1929.† It has been generally accepted and confirmed for a large variety of substances.

The figures for *para*-disubstituted benzene derivatives of the type  $X.C_6H_4.Y$  are of particular importance for two reasons. First, it was anticipated by Ingold‡ on the ground of the electronic theory of aromatic reactivity, that some deviation from the additive law might occur in this class of compound when

\* 'Phil. Mag.,' vol. 46, p. 513 (1923).

† 'Phys. Z.,' vol. 30, p. 397 (1929).

‡ 'Ann. Rep. Chem. Soc.,' vol. 23, p. 149 (1926).

Y was a group such as  $\text{—NR}_2$  as a result of electronic displacements in the aromatic nucleus. When the necessary data became available by the measurements of Højendahl,\* a gross discrepancy of this kind appeared probable for *p*-nitraniline and *p*-nitroanisole, and this was specifically emphasized at that time.† Secondly, a number of workers‡ have recently made use of dipole moment data for aromatic ethers and sulphides of this same general type in order to compute the angle between the valencies of the characteristic atom (oxygen or sulphur) of the group Y.

We have therefore carried out a direct analysis of the more numerous data now available for *p*-substituted anisoles, diphenyl ethers, phenols, and anilines in order to determine as far as possible the extent of this anomaly and its bearing on the problem of valency angles.

### *Basis of the Analysis.*

For the purpose of this analysis it is necessary to make the following postulates:—

(a) The observed dipole moments of the compounds discussed are strictly the vector sums of the constituent dipole moments within the molecule. This is an axiom upon which all analyses of dipole data must be based.

(b) The dipole moments of the compounds  $\text{C}_6\text{H}_5\text{X}$  may be regarded as giving the bond moments for the  $\text{C—X}$  bonds in aromatic substances, so that in general the moment of  $\text{X} \cdot \text{C}_6\text{H}_4 \cdot \text{Y}$  should be the vector sum of the moments of  $\text{C}_6\text{H}_5\text{X}$  and  $\text{C}_6\text{H}_5\text{Y}$ , except so far as they are modified as suggested under (f). In the absence of *ortho*-disubstituted compounds this convention is generally accepted.

(c) The moments of  $\text{O—CH}_3$ ,  $\text{O—H}$ , and  $\text{N—H}$  bonds in anisole, phenol, and aniline are constant and unaffected by nuclear substitution, among the compounds discussed. Any variation in the magnitudes of these bond moments which may occur is probably negligible.

(d) The valency angle  $\theta$  of oxygen remains constant in each group of substances—phenols, anisoles, or diaryl ethers—but the three values  $90^\circ$ ,  $109^\circ 28'$ ,

\* 'Phys. Z.', vol. 30, p. 391 (1929).

† Bennett, 'Ann. Rep. Chem. Soc.', vol. 26, p. 132 (1929).

‡ Bergmann, Engel, and Sándor, 'Z. phys. Chem.', B, vol. 10, p. 397 (1930); Smyth and Walls, 'J. Amer. Chem. Soc.', vol. 54, p. 3230 (1932); Hampson and Sutton, 'Proc. Roy. Soc.', A, vol. 140, p. 562 (1933); Hampson, Farmer, and Sutton, *ibid.*, A, vol. 143, p. 147 (1933).

and  $130^\circ$  will be taken as alternative possibilities throughout. Small variations in  $\theta$  are no doubt possible, but their order is probably such as to have little effect on the conclusions reached in this paper. The value of  $142^\circ \pm 8^\circ$  for  $\theta$ , recently suggested by Hampson, Farmer, and Sutton (*loc. cit.*), is rejected because evidence from the infra-red spectrum of water vapour,\* the collision areas of dimethyl and diphenyl ethers,† and the crystal structures of ice‡ and of arsenious and antimonious oxides§ makes it appear improbable. Further, the large angle  $\theta$  requires the O—Ph and O—CH<sub>3</sub> bond moments to have values which are markedly irregular when inserted in a table of the moments of such bonds for the elements arranged in the order of the periodic classification. Lastly, the proposed bond moment¶ for O—CH<sub>3</sub> is also inconsistent with the observed moment of ethylene oxide in which the valency angle is fixed by the geometry of the molecule.

(e) The valency angle of nitrogen in the amines is  $109^\circ 28'$ : this is a probable value and one very close to that computed for the ammonia molecule from spectral data.|| The possible angle  $90^\circ$  has been considered,¶¶ but rejected because it does not lead to consistent values of the CH<sub>3</sub>—N, H—N, and C<sub>6</sub>H<sub>5</sub>—N bond moments.

(f) Any difference between the observed values of the dipole moment of a substance X.C<sub>6</sub>H<sub>4</sub>.Y and the value computed by the vector summation is attributed to the existence of an additional moment  $\Delta\chi$  in the direction, positive or negative, of the axis of the benzene nucleus joining the *para*-positions occupied by the substituents X and Y.

### *p*-Substituted Anisoles.

From the known dipole moment of dimethyl ether ( $\mu = 1.29$  units: the unit adopted throughout is  $10^{-18}$  e.s.u.\*\*\*) and of diphenyl ether ( $\mu = 1.12$ ) the

\* Lueg and Hedfeld, 'Z. Physik,' vol. 75, p. 512 (1932); Plyler, 'Phys. Rev.,' vol. 39, p. 77 (1932); Mecke, Baumann, and Freudenberg, 'Z. Physik,' vol. 81, pp. 313, 445, 465 (1933). Cf. Van Vleck and Cross, 'J. Chem. Phys.,' vol. 1, p. 357 (1933).

† Hare and Mack, 'J. Amer. Chem. Soc.,' vol. 54, p. 4272 (1932).

‡ Barnes, 'Proc. Roy. Soc.,' A, vol. 125, p. 670 (1929).

§ Bozorth, 'J. Amer. Chem. Soc.,' vol. 45, p. 1621 (1923).

|| Dennison and Uhlenbeck, 'Phys. Rev.,' vol. 41, p. 313 (1932); Lueg and Hedfeld, 'Z. Physik,' vol. 75, p. 599 (1932).

¶¶ Cf. Pauling, 'J. Amer. Chem. Soc.,' vol. 53, p. 1367 (1931).

\*\*\* Data except where otherwise specified are taken from the lists in Freudenberg's "Stereochemie," 1932, or Smyth's "Dielectric Constant and Molecular Structure," 1931.

following bond moments are at once derived for the three values of  $\theta$  considered :—

	$\theta = 90^\circ$ .	$\theta = 109^\circ 28'$ .	$\theta = 130^\circ$ .
$m(\text{O}-\text{CH}_3)$ .....	0.91	1.12	1.53
$m(\text{O}-\text{C}_6\text{H}_5)$ .....	0.79	0.97	1.33

These values may be compounded to give a computed dipole moment for anisole in good agreement with that observed. (Calculated : 1.21 ( $\theta = 90^\circ$ ), 1.21 ( $\theta = 109.5^\circ$ ), 1.22 ( $\theta = 130^\circ$ ). Observed :  $\mu = 1.19$ .)

The dipole moments of the compounds  $\text{C}_6\text{H}_5\text{X}$  have been taken as follows :—

X	$\text{CH}_3$	F	Cl	Br	I	$\text{NO}_2$
$\mu$	0.4	1.42	1.55	1.52	1.30	3.90

In the calculation of  $\Delta_X^{\text{OMe}}$  it is found that two alternative values are mathematically possible; when these are considered, however, with reference to the direction and relative magnitude of the known electron-attracting or repelling property of the particular group X the chemically correct solution becomes obvious. When X is  $\text{CH}_3$  an independent confirmation for the choice made can be obtained from an examination of the data for the methyl cresyl ethers.

The results obtained for the six anisoles, using the three chosen values of  $\theta$ , are given in Table I.†

Table I.— $\text{Y} = \text{OCH}_3$ .

X.	$\mu_{\text{obs}}$	$\theta = 90^\circ$ .		$\theta = 109.5^\circ$ .		$\theta = 130^\circ$ .	
		$\mu_{\text{calc}}$	$\Delta_X^{\text{OMe}}$ .	$\mu_{\text{calc}}$	$\Delta_X^{\text{OMe}}$ .	$\mu_{\text{calc}}$	$\Delta_X^{\text{OMe}}$ .
$\text{CH}_3$ ..	1.20	1.50	+0.41	1.45	+0.43	1.39	+0.49
*F ..	2.09	1.11	+1.25	1.34	+0.98	1.59	+0.60
†Cl ..	2.24	1.19	+1.29	1.42	+1.02	1.68	+0.70
*Br ..	2.27	1.17	+1.35	1.40	+1.09	1.66	+0.77
*I ..	2.12	1.04	+1.40	1.27	+1.14	1.53	+0.78
‡ $\text{NO}_2$	4.74	3.24	+1.54	3.47	+1.32	3.72	+1.04

\* Bergmann and Tschudnowsky, 'Z. phys. Chem.,' B, vol. 17, pp. 100, 107 (1932).

† Bergmann and Engel, *ibid.*, B, vol. 15, p. 85 (1931).

‡ Donle and Gehrold, 'Z. phys. Chem.,' B, vol. 18, p. 316 (1932); Cowley and Partington, 'J. Chem. Soc.,' p. 1252 (1933).

The figures show that there is always a large additional moment  $\Delta_X^{\text{OMe}}$  operating in a direction from the oxygen atom across the nucleus, the value of which (a) is much larger than any possible error in the data; (b) depends

§ The arbitrary convention is adopted that the sign of  $\Delta$  is positive when it acts in the direction away from the oxygen atom.

somewhat on the value of the angle chosen, but the trend of the figures is always substantially the same; (c) varies with the nature of the atom or group X in the sequence methyl, halogens, nitro-group, which is the order of the dipole moments of the respective compounds  $C_6H_5X$ , and within the halogen family it follows a natural order from fluorine to iodine. A direct electrical interaction of the groups in the *para*-positions should be of negligible magnitude, and it is impossible to resist the conclusion that these figures are evidence of a special electronic redistribution in the system  $X-C_6H_4-O$ —due primarily to an effect emanating from the oxygen atom and varying with the electron-attracting or repelling properties of the substituent X.

#### p-Substituted Phenols.

Proceeding as before from the known dipole moments of diphenyl ether (1.12) and phenol (1.56) the bond moments found in phenol for the three values of  $\theta$  are:—

	$\theta = 90^\circ.$	$\theta = 109^\circ 28'.$	$\theta = 130^\circ.$
$m(O-C_6H_5)$ .....	0.79	0.97	1.33
$m(O-H)$ .....	1.34	1.58	2.04

Once again the available data\* are shown in Table II with the calculated dipole moments and the vector difference  $\Delta_X^{OH}$  along the axis for each substituent and each value of  $\theta$ .

Table II.—Y = OH.

X.	$\mu_{obs.}$	$\theta = 90^\circ.$		$\theta = 109.5^\circ.$		$\theta = 130^\circ.$	
		$\mu_{calc.}$	$\Delta_X^{OH}.$	$\mu_{calc.}$	$\Delta_X^{OH}.$	$\mu_{calc.}$	$\Delta_X^{OH}.$
$CH_3$ .	1.57	1.80	+0.38	1.71	+0.34	1.61	+0.25
Cl .	2.22	1.55	+1.01	1.85	+0.54	2.18	+0.05
Br .	2.13	1.53	+0.92	1.83	+0.45	2.16	-0.05
$NO_2$ . . .	5.02	3.39	+1.73	3.76	+1.34	4.18	+0.89

The values of  $\Delta_X^{OH}$  must be considered in relation to the corresponding figures in the anisole series, and it is apparent that a consistent conclusion from both sets of data can only be reached by assuming a value of  $\theta$  of less than  $130^\circ$ .

#### p-Substituted Diphenyl Ethers.

The requisite bond moments have been derived above. The figures for the diphenyl ethers are given in Table III.

\* Donle and Gehreken, *loc. cit.*



Table III.—Y = O.Ar.

Ether.	$\mu_{\text{obs}}$	$\theta = 90^\circ$ .		$\theta = 109.5^\circ$ .		$\theta = 130^\circ$ .	
		$\mu_{\text{calc.}}$	$\Delta_{\text{X}}^{\text{OAr}}$	$\mu_{\text{calc.}}$	$\Delta_{\text{X}}^{\text{OAr}}$	$\mu_{\text{calc.}}$	$\Delta_{\text{X}}^{\text{OAr}}$
*Ph. O. C <sub>6</sub> H <sub>4</sub> . CH <sub>3</sub>	1.31	1.43	+0.15	1.39	+0.11	1.34	+0.05
†Ph. O. C <sub>6</sub> H <sub>4</sub> . Br	1.57	1.08	+0.63	1.26	+0.40	1.46	+0.15
*Ph. O. C <sub>6</sub> H <sub>4</sub> . NO <sub>2</sub>	4.20	3.21	+1.01	3.38	+0.84	3.57	+0.65
†(Br. C <sub>6</sub> H <sub>4</sub> ) <sub>2</sub> O	0.60	1.03	-0.30†	0.63	-0.03†	0.16	+0.52†
†(NO <sub>2</sub> C <sub>6</sub> H <sub>4</sub> ) <sub>2</sub> O	2.80	4.40	-1.13†	3.38	-0.60†	2.17	+0.75†

\* Bergmann and Tschudnowsky, *loc. cit.*† Smyth and Walls, *loc. cit.*‡ Along *each* nuclear axis.

These figures again reveal the existence of considerable deviations  $\Delta_{\text{X}}^{\text{OAr}}$  among the diphenyl ethers from the simple vector additivity. The data are fewer, but the general conclusion may be drawn that the values of  $\Delta_{\text{X}}^{\text{OAr}}$  for the diphenyl ethers are smaller than those for the anisoles, and the values for ethers substituted in both nuclei are, taking  $\theta = 90^\circ$  or  $109.5^\circ$ , zero or of reversed sign. It may be concluded that the effect  $\Delta_{\text{X}}^{\text{OAr}}$  for an ether of the formula X.C<sub>6</sub>H<sub>4</sub>.OR varies with the nature of the group R(CH<sub>3</sub>, C<sub>6</sub>H<sub>5</sub>, X.C<sub>6</sub>H<sub>4</sub>). A diminution of the effect with such changes in R is consistent with our view that it represents a polarization or "mesomeric effect"§ which should be reduced by diminution in the electron availability at the oxygen atom; but the change of sign would be surprising, and suggests that the valency angle may in these substances be somewhat larger than  $109.5^\circ$ .

#### p-Substituted Anilines.

Assuming a valency angle of  $109^\circ 28'$  a set of consistent bond moments in the amine group has been obtained, which are given in Table IV below, together with the principal molecular moments on which they are based, the agreement between observed and calculated figures being shown.

On the basis of Table IV the data for substituted anilines are given in Table V.

Alternative mathematically possible values of  $\Delta_{\text{X}}^{\text{NH}_2}$  have again been rejected as unreal from the chemical point of view.

The  $\Delta_{\text{X}}^{\text{NH}_2}$  values in the above table are too large to be ignored and have somewhat the same kind of range as those for the anisoles: the figures for CH<sub>3</sub>, F, Cl, and Br are here, however, more nearly equal. The greater basic

§ Ingold, 'J. Chem. Soc.,' p. 1120 (1933).

function and directive power in nuclear substitution of the  $\text{NH}_2$  as compared with the OR group would lead us to expect that  $\Delta_{\text{X}}^{\text{NH}_2}$  would be larger than  $\Delta_{\text{X}}^{\text{OR}}$ . The figures show this to be true only for the group  $\text{NO}_2$ , and that for other substituents the values are nearly equal, yet somewhat smaller in the aniline series. This discrepancy is, however, at least in part accounted for when the bond moments  $\text{O}-\text{CH}_3$ ,  $\text{O}-\text{C}_6\text{H}_5$  and  $\text{N}-\text{CH}_3$ ,  $\text{N}-\text{C}_6\text{H}_5$  are compared. The  $\text{O}-\text{C}_6\text{H}_5$  moment in unsubstituted anisole differs but slightly

Table IV.

Bond Bond moment	$\text{CH}_3-\text{N}$ 0.71		$\text{H}-\text{N}$ 1.23	$\text{C}_6\text{H}_5-\text{N}^*$ 0.21		
Amine ...	$\text{MeNH}_2$	$\text{Me}_2\text{NH}$	$\text{Me}_3\text{N}$	$\text{PhNH}_2$	$\text{Ph}_2\text{NH}$	$\text{Ph}_3\text{N}^\dagger$
$\mu_{\text{calc}}$	(1.23)	1.01	0.71	(1.55)	1.32	0.21
$\mu_{\text{obs}}$	1.23	0.96	0.6	1.55	1.3	0.26

\* The figure 0.4 for the bond moment of  $\text{Ph}-\text{N}$  has been used by Sidgwick ('Chem. Rev.', vol. 9, p. 77 (1931)). If this is adopted in place of our value, the agreement among the amines is less satisfactory, but the general conclusions concerning  $\Delta$  remain the same. The value 1.3 for this bond (Hammick, New, Sidgwick, and Sutton, 'J. Chem. Soc.', p. 1876 (1930); Sutton and Thomas, *ibid.*, p. 406 (1933)) is less acceptable.

† Bergmann and Schütz, 'Z. phys. Chem.', B, vol. 19, p. 401 (1932).

Table V.  $-\text{Y} = \text{NH}_2$ .

X.	$\mu_{\text{obs}}$	$\mu_{\text{calc}}$	$\Delta_{\text{X}}^{\text{NH}_2}$
* $\text{CH}_3$	1.27	1.18	+0.31
†F	2.75	2.36	+0.43
Cl	2.90	2.40	+0.49
Br	2.87	2.42	+0.50
I	2.82	2.23	+0.66
† $\text{NO}_2$	6.20	4.06	+1.58

\* Donle and Gehreckens, *loc. cit.*

† Bergmann and Tschudnowsky, *loc. cit.*

from the  $\text{O}-\text{CH}_3$  moment (0.97 and 1.12 if  $\theta = 109.5^\circ$ ) whereas the  $\text{N}-\text{C}_6\text{H}_5$  value (0.21) is already much smaller than the  $\text{N}-\text{CH}_3$  figure (0.71) so that part of the polarization which arises among the anilines is present in the parent substance and the computed  $\Delta_{\text{X}}^{\text{NH}_2}$  values are proportionately reduced.

### *Criticism of the Method for Determining the Valency Angles of Oxygen and Sulphur by means of Dipole Moment Data.*

The method referred to is based on the fact that the change of dipole moment resulting from the introduction of a known substituent X into the *para*-position is determined by the angle between the axis in which this substituent

lies and the direction of the moment of the parent compound into which it has been introduced. The validity of this method depends on the accuracy with which the resultant moment of the molecule may be obtained by vectorial addition of the values for the substituent dipoles taken from the monosubstituted benzene compounds ( $C_6H_5 \cdot X$ ). In terms of the preceding analysis this requires that  $\Delta_X^Y = 0$ , which has been shown to be far from true in each of the series of anisoles, phenols, diphenyl ethers, and anilines.

The method has nevertheless been advocated in recent years by Bergmann and his colleagues, by Smyth and Walls, and by Hampson, Farmer, and Sutton (*loc. cit.*). Of these authors Bergmann regards the method as inexact when using the diphenyl oxides and sulphides; Smyth admits the possibility of error from interaction of the groups, but considers it negligible; Hampson and Sutton,\* after a critical examination of the point, conclude that there is no serious error involved except "when the two groups cause large and opposed electromeric effects." It becomes necessary to examine the arguments of these workers in order to account for the difference in our conclusions.

Bergmann and Sutton have approached the problem by a method less direct than our own analysis above. The angle between the resultant dipole moment of the compound  $C_6H_5Y$  and the axis of the benzene ring, being unknown, is designated  $\phi$ ,† and is calculated or obtained graphically from the vector triangle formed from the known dipole moments of the three compounds  $C_6H_5X$ ,  $C_6H_5Y$ , and  $X \cdot C_6H_5$ . The validity of the assumption (made for the purpose of this calculation) that there is no deviation from exact additivity is then regarded as confirmed if the values of  $\phi$  found for any one group Y are constant, within certain limits, for several different substituents X. The computed value of  $\phi$  was found to be constant for the phenols by Bergmann, Engel, and Sándor (*loc. cit.*), and the same conclusion has been reached more recently by Donle and Gehrckens (*loc. cit.*). A similar constancy was found for  $\phi$  among substituted anisoles.‡ Hampson and Sutton have recalculated the results of Bergmann and claim that  $\phi$  varies only from  $71.5^\circ$  to  $80^\circ$  in the series of halogeno-, nitro-, and methyl-anisoles. The figure  $48^\circ$  calculated for iodo-anisole is rejected as probably incorrect. It should, however, be pointed out that owing to an error of transcription made in the paper of Bergmann and Tschudnowsky, and copied by Hampson and Sutton, the data

\* *Loc. cit.*, p. 573.

† Bergmann and Engel, 'Z. phys. Chem.,' B, vol. 8, p. 110 (1930); Bergmann, Engel, and Sándor, *loc. cit.*; cf. Højendahl, *loc. cit.*

‡ Bergmann and Tschudnowsky, *loc. cit.*

in the paper by the latter authors for bromo- and iodo-anisole, which had become interchanged, require correction. The values of  $\phi$  have, therefore, again been recalculated using the correct moments for these substances and the newer value  $\mu = 4.74$  for *p*-nitroanisole. The following are the figures :—

Substituent X .....	CH <sub>3</sub>	F	Cl	Br	I	NO <sub>2</sub>
Value of $\phi$ for <i>p</i> -X.C <sub>6</sub> H <sub>4</sub> .OMe..	82°	74°	71°	67°	63°	51°

The following points require consideration :—

(1) These values of  $\phi$  are not constant, and their variation is evidently systematic since it follows the polar sequence methyl : halogens : nitro-group, and the order of the atomic numbers within the halogen family.

(2) The figure of 82° for the methyl group is particularly liable to error because the vector triangle is, in this instance, of such a shape that small differences in the data used influence the calculated value of  $\phi$  to a disproportionate extent.

(3) All the authors who have computed the angle  $\phi$  appear to have overlooked the fact that constancy of  $\phi$  would be no evidence that the additivity was perfect and the value of the angle obtained a true one. The condition that  $\mu_{PhX}$  and  $\mu_{PhY}$  on the one hand and  $(\mu_{PhX} + \Delta_X)$  and  $\mu_{PhY}$  on the other should both sum vectorially to the same value  $\mu_X C_6H_4 Y$  is given by the equation

$$\Delta_X^2 + 2\Delta_X (\mu_{PhX} + \mu_{PhY} \cos \beta') + 2\mu_{PhX} \mu_{PhY} (\cos \beta' - \cos \beta) = 0,$$

where  $\beta$  and  $\beta'$  are the angles between the directions of  $\mu_{PhY}$  and  $\mu_{PhX}$  in the two cases considered—that is, the respective values which will be found for the angle  $\phi$ . It would thus be mathematically possible for a suitable series of deviations  $\Delta_X$  to cause a false yet constant value of  $\phi$  to appear throughout.

The general conclusion is drawn that the mutual action of groups leads to a serious and variable deviation from simple vector additivity of dipole moments in the phenols, anisoles, diphenyl ethers, and anilines, and that in consequence it is unsafe to base computations of valency angles on such data. This must be held to apply also to the sulphides.

### Summary.

The data for the dipole moments of *para*-substituted anisoles, phenols, diphenyl ethers, and anilines have been analysed by a direct method. The conclusion is reached that in all these substances a deviation  $\Delta$  from strict additivity of the bond moments occurs. The values of  $\Delta$ , which is an addi-

tional moment operating along the axis of the benzene nucleus, are consistent with those to be expected from the point of view of the electronic theory of aromatic reactivity, and may be regarded as due to the "mesomeric" effect in such substances.

The use which has been made of the dipole moments of aromatic compounds for the computation of the valency angle of oxygen (and by implication of sulphur) is shown to be unsound.

---

### *The Nature and Properties of Aluminosilicate Framework Structures.*

By W. H. TAYLOR, Manchester University.

(Communicated by W. L. Bragg, F.R.S.—Received December 21, 1933.)

#### I. INTRODUCTION.

In recent years it has been shown by X-ray methods that the structures of a large number of crystals are based on frameworks of linked tetrahedral groups of oxygen atoms. The individual tetrahedra contain silicon or aluminium atoms, and other atoms (such as sodium or potassium) and water molecules or molecular groups (such as  $\text{CO}_3$  and  $\text{SO}_4$ ) are located in the interstices of the oxygen atom arrangement. Some of the structure determinations are incomplete and lack direct experimental proof, but in others it has been possible to discover the details of the atomic arrangement. The present writer has been associated with several of these detailed investigations, and in this paper presents some general conclusions which may be drawn from an examination of the available data. Experimental details and evidence for the correctness of individual structures have been published elsewhere.

For the purpose of the present paper a framework structure is defined as one in which every tetrahedron  $\text{SiO}_4$  or  $\text{AlO}_4$  shares all its corners with other tetrahedra, thus accounting for all the silicon aluminium and oxygen atoms in the structure; such a crystal has a chemical formula in which the ratio  $(\text{Si} + \text{Al})$  to  $\text{O}$  is 1 to 2. Framework structures include the forms of silica, the feldspars, the zeolites, the ultramarines, nepheline and kaliophilite, and related compounds which will be mentioned later; also danburite if boron may be supposed to take the place of aluminium in our definition, and probably leucite.

The following sections of this paper deal with the various types of framework structure, the nature of the kation environment, the water-content of zeolites, and some aspects of the problems of twinning and pseudo-symmetry of framework structures in general.

## II. TYPES OF FRAMEWORK STRUCTURE.

Tetrahedra of oxygen atoms, with silicon or aluminium atoms at the centres, are linked by sharing all corners and so build a strong framework which forms the basis of the structure. The framework may be either infinitely extended in all directions in space, or infinitely extended in two directions and of finite thickness, and may be either rigid and self-supporting or collapsible. If it is collapsible, the equilibrium configuration is determined by the kations and water molecules (or other molecular groups) located in the interstices of the framework. Examples of the various types of framework are described in this section.

In all framework structures the fundamental unit is the tetrahedron  $\text{SiO}_4$  or  $\text{AlO}_4$ , but another unit which occurs in all is the ring of four tetrahedra one form of which is illustrated in fig. 2. The tetrahedra in the ring are tilted in different ways according to the nature of the links with adjacent tetrahedra, but the ring itself is apparently a stable unit. The ring of six tetrahedra illustrated in fig. 3 also occurs frequently, but it is often much distorted from the ideal form illustrated; it is probably better regarded as an incidental consequence of the linkage of four-rings than as a definite unit of structure. (In mica\* the basis of the structure is a sheet of linked six-rings almost undistorted from the ideal configuration. This structure, however, is not of the framework type in the terms of our definition.)

### A. Three-dimensional Frameworks.

We consider now frameworks which are infinitely extended in all directions, and so form three-dimensionally infinite anions. To this class belong all framework structures hitherto described. The various structures may be classified according to the rigidity of the framework.

(i) *Rigid and Self-supporting Frameworks.* --It is not quite certain that any completely rigid framework structures are known, but there is some evidence that analcite and chabazite may belong to this group.

\* Jackson and West, 'Z. Kristallog.' vol. 76, p. 211 (1930).

The zeolite analcite\* has cubic or pseudocubic symmetry, and the unit cell of edge  $a = 13.68 \pm 0.04$  Å. contains 16 molecules of composition  $\text{NaAlSi}_2\text{O}_6 \cdot \text{H}_2\text{O}$ . No departure from cubic symmetry has been detected in X-ray photographs, and corresponding to this the 16  $\text{AlO}_4$  groups in the framework cannot be distinguished from the 32  $\text{SiO}_4$  groups. Further, it is necessary in order to explain the X-ray intensities to assume that the 16 sodium atoms occupy (uniformly on the average) 24 equivalent positions in the cubic unit cell (see also Section V). To test the correctness of this assumption silver analcite ( $\text{AgAlSi}_2\text{O}_6 \cdot \text{H}_2\text{O}$ ) was prepared and powder photographs of ordinary sodium analcite and silver analcite were compared. The alterations in relative intensities prove that the sodium atoms are correctly placed in the structure, and correspond to fairly complete (say 70%) replacement of sodium by silver. The unit-cell dimensions, however, are unchanged within the limits of experimental error ( $\pm 0.04$  Å.), although the silver ion (radius 1.26 Å. according to Pauling†) is definitely larger than the sodium ion (radius 0.98 Å.) which it replaces. If this result is confirmed by measurements on chemically-analysed material, it must indicate that the tetrahedra can suffer slight re-orientation locally in the framework in order to accommodate the larger ion, without altering the shape and size of the framework as a whole. In this sense the analcite framework is rigid.

Chabazite is one of the most highly hydrated zeolites. Its unit cell and space group have been determined by Wyart,‡ its structure is being investigated in this laboratory at the present time. The work is incomplete, but it appears that the framework cannot be distorted appreciably without departing from the observed crystal symmetry, and that it is probable that the kations occupy positions adjacent to large cavities which also contain the water molecules. If these suggestions prove to be correct, little alteration in the dimensions of the framework should follow the replacement of calcium by larger ions. So far as I am aware, no data is yet available to test this hypothesis.§

(ii) *Frameworks containing Rigid Units.*—The structures of the fibrous zeolites edingtonite,|| thomsonite, natrolite, scolecite, mesolite¶ are closely related. The string or chain of tetrahedra illustrated in fig. 1 (*d'*) is the unit

\* Taylor, 'Z. Kristallog.,' vol. 74, p. 1 (1930).

† 'Z. Kristallog.,' vol. 67, p. 377 (1928).

‡ 'C. R. Acad. Sci. Paris,' vol. 192, p. 1244 (1931).

§ Wyart has described the structure of chabazite ('Bull. Soc. franc. Minér.,' vol. 56, p. 81 (1933)). His results confirm the suggestions made above.

|| Taylor and Jackson, 'Z. Kristallog.,' vol. 86, p. 53 (1933).

¶ Taylor, Meek and Jackson, 'Z. Kristallog.,' vol. 84, p. 373 (1933).

common to all, the various structures being obtained by linking such chains in different ways.

The chain itself is a rigid structure. It is formed of linked four-rings by tilting the tetrahedron vertices in such a way as to form another tetrahedron between one four-ring and the next. The length of the chain-unit is 6.6 Å. and it includes five silicon or aluminium atoms, so that all crystals built from such chains must have a unit cell with one edge 6.6 Å. and containing 10 or 20 or 40 tetrahedra.

In the complete structure the tetrahedron chains lie side by side, and adjacent chains are linked by sharing oxygen atoms which belong to only one tetrahedron in each chain and are at the sides of the chains. The linking oxygen atom may lie either on a reflection plane or on a rotation axis; if all linking atoms lie on reflection planes the edingtonite framework results, if half lie on planes and half on rotation-axes the framework is that of thomsonite, and if all lie on rotation-axes the framework is that of natrolite scolecite and mesolite (neglecting the small departure from tetragonal symmetry in these last crystals). Figs. 1 (a), (b), (c) show the unit cells diagrammatically, and in Table I the dimensions are quoted. As a direct consequence of the nature of the frameworks, all the crystals are fibrous in growth in the direction of the *c*-axis,—

Table I.—Axial Lengths in the Fibrous Zeolites.

Compound.	Radius of kation.	<i>a</i> .	<i>b</i> .	<i>c</i> .
	Å.	Å.	Å.	Å.
BaAl <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 4H <sub>2</sub> O	1.4	9.6	9.7	6.5
Tl <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 2H <sub>2</sub> O	1.7 ?	9.9	10.0	6.65
K <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 2H <sub>2</sub> O	1.3	9.7	10.0	6.6
Ag <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub>	1.2	9.4	9.4	6.5
Na <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 2H <sub>2</sub> O	1.0	19.0	19.0	11.9
NaCa <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 6H <sub>2</sub> O	1.0	13.1	13.1	6.6
Na <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 2H <sub>2</sub> O	1.0	18.3	18.6	6.6
Li <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 2H <sub>2</sub> O	0.7	17.8	18.6	6.5
Ag <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 2H <sub>2</sub> O	1.2	18.6	18.9	6.6
(NH <sub>4</sub> ) <sub>2</sub> Al <sub>2</sub> Si <sub>2</sub> O <sub>10</sub>	1.4	17.9	18.4	6.6
CaAl <sub>2</sub> Si <sub>2</sub> O <sub>10</sub> · 3H <sub>2</sub> O	1.0	18.3	18.6	6.6

NOTES.

**Edingtonite.**—The barium compound is the natural mineral. The thallium and potassium substitution products are crystals of good quality, silver edingtonite is largely amorphous and gives very few weak X-ray reflections, sodium edingtonite gives X-ray photographs indicating that it is entirely different from the other edingtonites which all show general resemblance (hence the large alteration in the *c*-axis of this compound). The water contents given correspond approximately to the analyses by Hey. The axial lengths quoted for potassium edingtonite are pseudo-axes, the true axes are twice the lengths given.

**Thomsonite.**—The *c*-axial length quoted is a pseudo-axis, the true axis is twice the length given.

**Natrolite.**—The ammonium compound is anhydrous, and the loss of water apparently more than balances the effect of the bigger kation in determining the axes *a* and *b*. All the substituted natrolites are crystals of good quality.



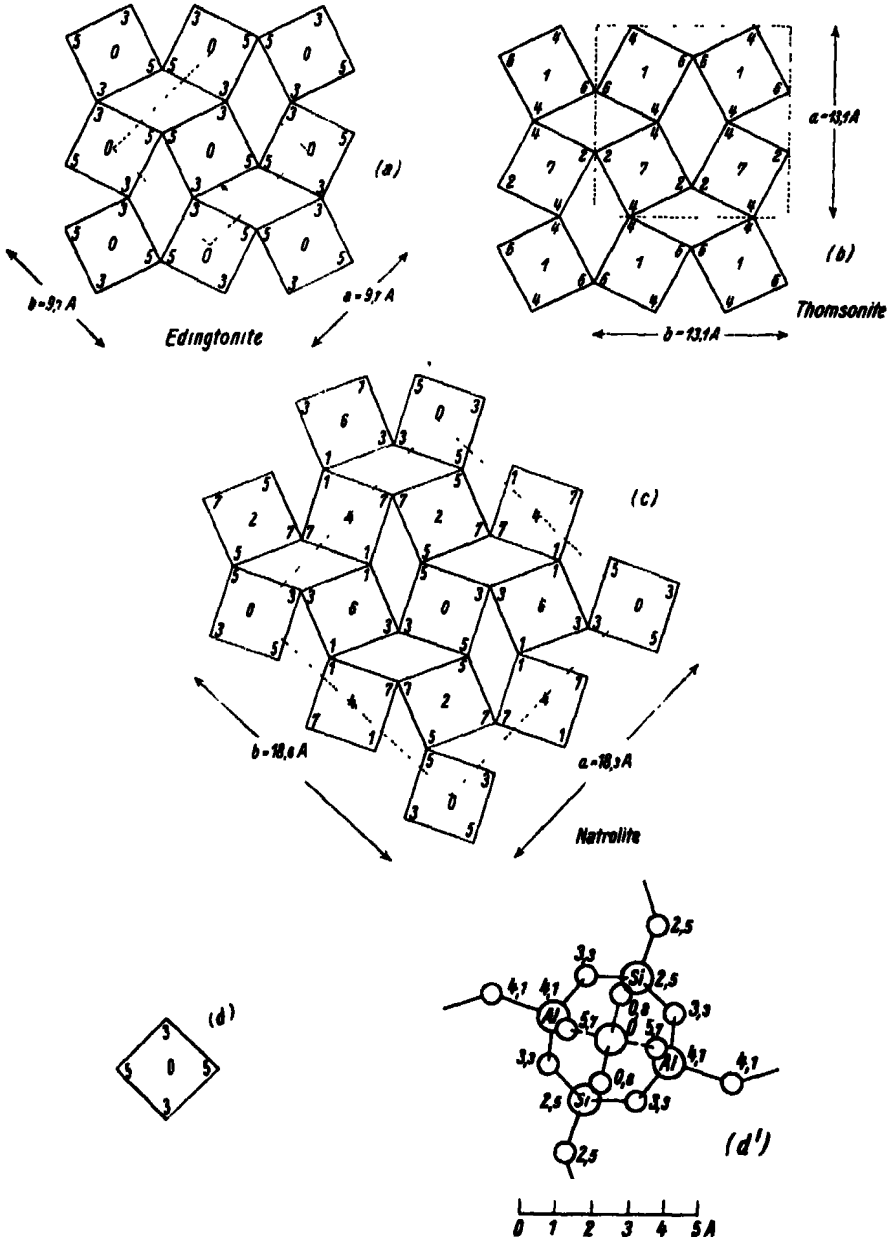


FIG. 1.—Linkage of tetrahedron chains in the fibrous zeolites. The structure of one tetrahedron chain is shown in (d'), in which the large circles represent silicon or aluminium atoms, the small circles represent oxygen atoms, and the heights of the atoms are given in Å. The same chain is represented diagrammatically in (d), where the numbers show the heights of silicon and aluminium atoms as multiples of  $c/8$  (the  $c$ -axis is  $6.6\text{ Å}$ , so that  $3c/8 = 2.5\text{ Å}$ ,  $5c/8 = 4.1\text{ Å}$ ). The linked chains are shown in (a), (b), (c) as arranged in edingtonite, thomsonite and the natrolite group, respectively, and in each the unit cell is indicated by dotted lines.

*i.e.*, parallel to the chain length—and all show right-angled cleavages (in edingtonite and natrolite {110}, in thomsonite (010) and (100)) which break the links between chains without breaking the chains themselves.

The kations and water molecules occupy positions in the channels between tetrahedron chains (see also Section IV and fig. 4), and the orientation of the chains is determined by the size of the kation. If the kation is large the chains are held widely apart and so the *a*- and *b*-axes are relatively long, if the kation is small the chains come more closely together and the *a*- and *b*-axes are shorter. Table I contains data which support this hypothesis regarding the equilibrium of the framework. Attention is directed to the following points.

In the first place the *c*-axis (*i.e.*, the chain length) is nearly the same in all the crystals, with the single exception of sodium edingtonite which Hey and Bannister (private communication) declare to be quite different in structure from the other edingtonites.

Secondly, Table I and figs. 1 (*a*), (*b*), (*c*) show that in the natural crystals barium edingtonite, thomsonite, sodium natrolite, and scolecite, the cell dimensions (in the plane 001) which should be compared are for edingtonite, *2a*, *2b*, *i.e.*, 19.2, 19.4 Å.; for thomsonite [110], [1 $\bar{1}$ 0], *i.e.*, 18.5, 18.5 Å.; for natrolite *a*, *b*, *i.e.*, 18.3, 18.6 Å.; for scolecite *a*, *b*, *i.e.*, 18.3, 18.6 Å. Thus although the kations in the three types of structures do not occupy precisely corresponding positions, corresponding lattice directions reflect the different sizes of the kations barium (in edingtonite) and sodium or calcium (in thomsonite, natrolite, and scolecite).

Lastly, it has been shown\* that the silver ions certainly, and the lithium ions very probably, occupy the positions of the sodium ions which they replace in the substituted natrolites, and these compounds show the expected alterations in the axial lengths *a* and *b*. Similarly in the edingtonite structures—the barium ions in natural edingtonite occupy one-half of the total number of available cavities, and it is highly probable that the thallium and potassium ions in the substitution products occupy all these cavities.

(iii) *Non-rigid Frameworks.*—Most of the framework structures are probably included in this group, important examples being the feldspars, the ultramarines and related crystals, and nepheline.

The feldspar framework† is built up from chains of four-rings of tetrahedra, illustrated in fig. 2, the length of the chain being parallel to the crystallographic

\* Taylor, Meek and Jackson, 'Z. Kristallog.', vol. 84, p. 373 (1933).

† Taylor, 'Z. Kristallog.', vol. 85, p. 425 (1933).

*a*-axis. Adjacent chains are linked to form the complete framework by sharing oxygen atoms which belong to only one tetrahedron in each chain. In the monoclinic feldspars the linking oxygen atoms lie either on reflection planes or

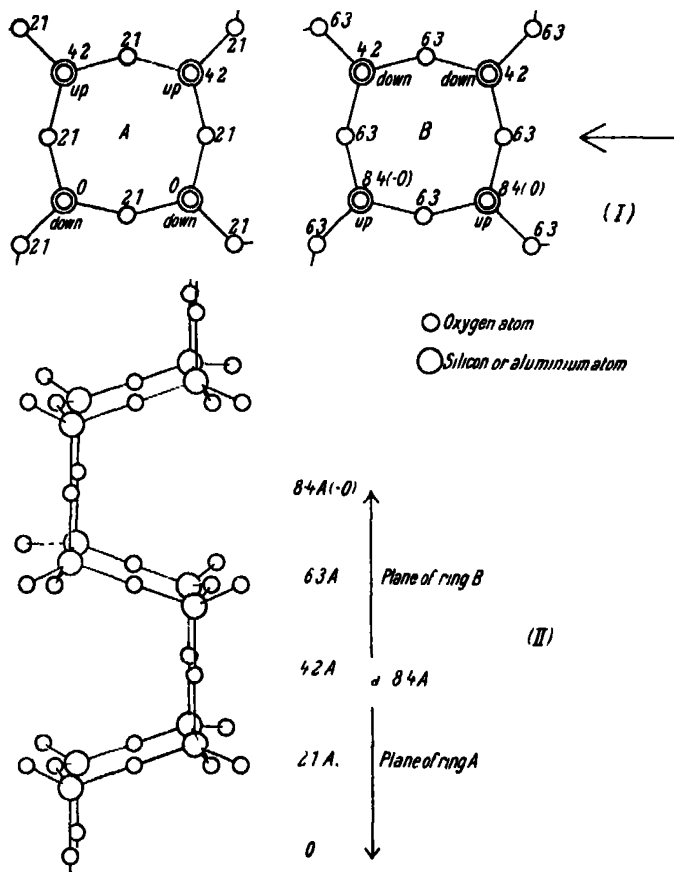


FIG. 2.—Idealized diagram of the chain of tetrahedra in the feldspars. In I the four-rings are viewed along the crystallographic *a*-axis, the numbers indicating the heights of oxygen atoms in A, above the base plane. Ring B is to lie above ring A in such a way that the oxygen atoms at height 4.2 Å. are common to both rings, and the four tetrahedra thus linked together (two from A, two from B) form a new ring of four. In II the linked rings are viewed in the direction indicated by the arrow in the upper diagram.

on rotation axes, and in the triclinic feldspars the atomic positions are all altered to a slight extent but the framework remains in essentials the same. The cleavages (001) and (010) break bonds between chains but do not break the strongly bound four-rings in the individual chains.

The X-ray examination of the feldspars recently made in this laboratory has shown that the equilibrium configuration of the framework is determined by the size of the kations which occupy certain cavities within the framework. In Table II the axial lengths and angles, the cleavage angle, and the radius of the kation, are quoted for the principal feldspars. Orthoclase, hyalophane, and celsian are monoclinic, albite and anorthite triclinic, and the division between the two symmetry-groups is a consequence of the fact that potassium and barium ions are large while sodium and calcium ions are small. Comparison

Table II.—Axial Lengths and Angles of the Feldspars.

Feldspar.	a.	b.	c.	$\alpha$ .		$\beta$ .	$\gamma$ .		C.	R.
	A.	A.	A.	°	'	°	'	°	'	A.
Orthoclase	8.4	12.9	7.1	90	0	116	3	90	0	1.3
Hyalophane	8.5	13.0	7.1	90	0	115	35	90	0	1.3-1.4
Celsian	8.6	13.1	7.3	(90	0)	(116	0)	(90	0)	1.4
Albite	8.1	12.0	7.1	94	3	116	29	88	9	1.0
Anorthite	8.2	12.95	14.2	93	13	115	55	91	12	1.0

*a, b, c* are the axial lengths,  $\alpha, \beta, \gamma$  the axial angles, C the cleavage angle, and R the radius of the kation.

The cleavage angle for celsian is almost exactly  $90^\circ$ , and the X-ray photographs correspond to the approximate values  $\alpha\beta\gamma$  given above.

In anorthite the period 7.1 Å. is a pseudo-axis, the true c-axis being 14.2 Å.

of X-ray intensities has shown\* that the barium ions in celsian ( $\text{BaAl}_2\text{Si}_2\text{O}_8$ ) occupy exactly the same positions as the potassium ions in orthoclase ( $\text{KAlSi}_3\text{O}_8$ ), and that in hyalophane there is a random replacement of potassium ions by barium ions according to the relative amounts of potassium and barium present; the framework is the same in all three crystals, only small dimensional differences being observed. In these crystals the orientation of the chains is such as to give a cleavage angle of  $90^\circ$ , and the individual chains are rather widely opened so that the *a*-axis (i.e., the chain length) is approximately 8.5 Å. In the plagioclases on the other hand the kations are the small sodium or calcium ions, and both the orientation of the chains, and the extent to which the individual chain is held open, are different. The changed orientation results in a cleavage angle of  $86^\circ$  approximately instead of  $90^\circ$ , and the closing-up of the individual chain shortens the *a*-axis (chain length) from 8.5 Å. to 8.2 Å.

It is worthy of remark at this point that the plagioclase feldspars do not form a single isomorphous series from albite ( $\text{NaAlSi}_3\text{O}_8$ ) to anorthite

\* Taylor, Darbyshire and Strunz, 'Z. Kristallog.' (*in press*).

( $\text{CaAl}_2\text{Si}_2\text{O}_8$ ), with a simple replacement of (NaSi) by (CaAl) which would correspond to the replacement of (KSi) by (BaAl) in the monoclinic feldspars. This fact is discussed in Section III; it does not invalidate the present discussion of framework equilibrium.

The ultramarines were the first framework structures to be described. Jaeger\* in 1929 showed that the cubic unit cell of an ultramarine is based on a framework of composition  $\text{Al}_3\text{Si}_3\text{O}_{24}$ , with the tetrahedra arranged in such a way as to form large cavities at the centre and corners of the unit cell. Within the cavities are placed the kations and molecular groups, some in fixed positions, some "wandering," according to Jaeger. Pauling afterwards showed that the structure of sodalite† ( $\text{NaCl} \cdot \text{Na}_3\text{Al}_3\text{Si}_3\text{O}_{12}$ ) is based on the same framework of linked tetrahedra. The framework is collapsible, and the equilibrium configuration is determined by the sizes of the extra kations and molecular groups. Thus the cubic unit cell edge has a maximum value  $a = 9.4$  Å. when the large alkali ions K, Rb, Cs, are introduced,‡ and the value  $a = 9.15$  Å. in hauynite ( $\text{CaSO}_4 \cdot \text{Na}_3\text{Al}_3\text{Si}_3\text{O}_{12}$ ) and noselite ( $\text{Na}_2\text{SO}_4 \cdot \text{Na}_3\text{Al}_3\text{Si}_3\text{O}_{12}$ ) which contain the large  $\text{SO}_4$  group. In sodalite the edge is  $a = 8.87$  Å. and in helvite ( $\text{MnS} \cdot \text{Mn}_3\text{Be}_3\text{Si}_3\text{O}_{12}$ ) it is  $a = 8.25$  Å., showing that the framework has collapsed very considerably from its most widely opened form because the sodium, manganese, chlorine, and sulphur ions are all relatively small. This group of crystals is quite remarkable for the extent to which the configuration of the framework may alter without any essential change in its structure.

The structures of nepheline and kaliophilite have not been fully worked out, but Bannister's measurements§ indicate that the cell of kaliophilite ( $\text{KAlSiO}_4$ ) is related in a very simple way to the nepheline cell if this is supposed to be expanded from the normal  $a = 9.96$  Å.,  $c = 8.35$  Å., for composition  $\text{NaAlSiO}_4$  to  $a = 10.4$  Å.,  $c = 8.5$  Å., when the sodium is almost completely replaced by the larger potassium ion.

### B. Frameworks of Finite Thickness.

No structure has yet been described in which all tetrahedron corners are shared in such a way as to form a framework infinite in two directions but of finite thickness. It has long been realized that such frameworks are geo-

\* 'Trans. Faraday Soc.,' vol. 25, p. 320 (1929).

† 'Z. Kristallog.,' vol. 74, p. 213 (1930).

‡ Jaeger and van Melle, 'Proc. Acad. Sci. Amst.,' vol. 30, pp. 479, 885 (1927).

§ 'Min. Mag.,' vol. 22, p. 569 (1931).

metrically possible, and there is reason to believe that the zeolites heulandite and stilbite are of this type.

An ideally simple example of such a tetrahedron framework is shown in fig. 3. First an infinite sheet of tetrahedra is formed by linking rings of six tetrahedra as in the mica structure, and all the tetrahedron vertices are supposed to point upward; if these vertices lie on a reflection plane of symmetry

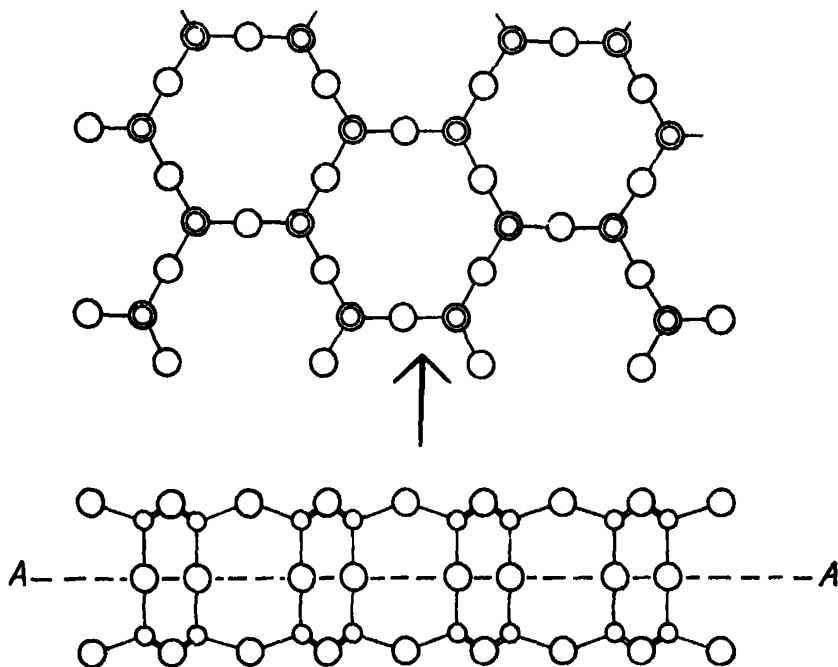


FIG. 3.—A tetrahedron framework of finite thickness. Large circles represent oxygen atoms, small circles silicon or aluminium atoms. The upper diagram represents a portion of an infinite sheet of tetrahedra in which all vertices are supposed to point upward. If these vertices lie on a reflection plane a second similar tetrahedron sheet, in which all vertices point down, is linked to the first. The lower diagram represents the appearance of the linked sheets, which form a tetrahedron framework of finite thickness, when viewed in the direction indicated by the arrow in the upper diagram.

they will link the first infinite sheet to a second similar sheet above the reflection plane, and there will result an infinite double sheet of finite thickness in which every tetrahedron shares all its corners with other tetrahedra.

The structure of heulandite is under examination at the present time in this laboratory. On the basis of the unit cell dimensions and space group first determined by Wyart\* we have failed to find any possible way of constructing

\* 'C. R. Acad. Sci. Paris,' vol. 190, p. 1564 (1930).

a continuous three-dimensional framework of linked tetrahedra. It is possible, however, to construct a sheet framework of finite thickness (more complicated than that illustrated in fig. 3) in conformity with the space group symmetry and unit cell dimensions. The plane of the sheets is parallel to (010), and successive sheets must be bound together by water molecules and kations which cannot be accommodated within the sheets. The investigation is not yet completed, but it is clear that such a structure of strong sheets bound together by relatively weak forces would explain both the highly perfect single cleavage (010) and also the interesting behaviour of the crystal on dehydration, studied by Wyart.\* He dehydrated heulandite gradually by heating in dried air, taking X-ray photographs at various stages, with results which may be summarized as follows. Heating from room temperature to  $210^{\circ}$  C. removed 10% of water and caused the length of the  $b$ -axis to fall from 17.8 Å. to 17.1 Å.; further heating resulted in marked changes in relative intensities of reflection, and in the appearance of asterism showing that the crystal was breaking down into a sub-parallel aggregate; at the same time the axial lengths (7.45, 17.8, 15.8 Å. at room temperature) fell to 7.26, 16.6, 15.2 Å., respectively, and on complete dehydration the length of the  $b$ -axis fell to 16.0 Å. The  $b$ -axis [010] is perpendicular to the plane of the framework sheets, and the removal of the water molecules which bind successive sheets together would be expected to produce a large alteration in its length, as observed by Wyart.

The unit cell and space group of stilbite (another zeolite showing a single highly perfect cleavage (010)) have been determined in this laboratory, and it is very significant that a similar sheet framework can be devised in accordance with the dimensions and symmetry of the crystal. We hope shortly to complete the investigation of these interesting structures, but I wish to emphasize that until this is done the above discussion of sheet frameworks must be regarded as speculative.

### III. KATIONS IN FRAMEWORK STRUCTURES.

The form of the framework of tetrahedra has an important bearing on the nature and extent of the isomorphous replacement which the included kations may suffer. For in the first place it determines the nature of the group of oxygen atoms around the kation, and in the second place it decides whether base-exchange (*i.e.*, artificially-promoted isomorphous replacement) can occur in the actual crystal specimen. Water molecules also influence base-exchange, and so, in all probability, does the charge on the kation.

\* 'C. R. Acad. Sci. Paris,' vol. 191, p. 1343 (1930).

The larger the kation, the less regular (in general) is the group of oxygen atoms around it, and the more difficult is it to decide which of the neighbouring atoms belong to the kation and which do not. The following is a summary of the nature of the groups around sodium, calcium, potassium, and barium ions in such framework structures as have been completely determined.

The sodium ion is surrounded by four oxygen atoms and two water molecules at 2.4 Å. in analcite\* ; by four oxygen atoms and two water molecules at 2.5 Å. (average distance) in natrolite,† and by six oxygen atoms at 2.6 Å. (average distance) in albite.‡ According to Pauling each sodium atom in sodalite§ is surrounded by a group of four anions--one chlorine ion and three oxygen ions.

The environment of the calcium ion is less certain. In scolecite|| it is probably surrounded by four oxygen atoms and three water molecules at 2.5 Å. approximately, and in thomsonite|| one-half of the calcium ions are probably surrounded by similar seven-fold groups, the remainder by an eight-fold group consisting of two oxygen atoms and two water molecules at 2.3 Å., with four more oxygen atoms at 3.0 Å. The environment of the calcium ion in anorthite must be different from that of the sodium ion in albite, and there is some reason for suggesting¶ that the group is seven-fold ; on account of the complexity of the structure the nature of the group cannot be determined by direct calculation.

The potassium and barium ions in orthoclase, hyalophane and celsian¶ are surrounded by six oxygen atoms at 2.85 Å. and four at 3.1 Å. (approximate distances), and in edingtonite\*\* there is a similar ten-fold group of oxygen atoms and water molecules around the barium ion.

In minerals with framework structures two types of isomorphous replacement are observed. In the first the exchange is represented by  $\text{KSi} \rightleftharpoons \text{BaAl}$  or  $\text{NaSi} \rightleftharpoons \text{CaAl}$ , and the total number of kations is unaltered, valencies being balanced by an equivalent alteration in the aluminium-to-silicon ratio. The second type is represented by  $\text{Ba} \rightleftharpoons \text{K}_2$  or  $\text{Ca} \rightleftharpoons \text{Na}_2$ , and in this type there is no alteration in the amount of aluminium but the total number of kations varies.

\* Taylor, 'Z. Kristallog.' vol. 74, p. 1 (1930).

† Taylor, Meek and Jackson, *ibid.*, vol. 84, p. 373 (1933).

‡ Taylor, Darbyshire and Strunz, *ibid.* (*in press*).

§ Pauling, *ibid.*, vol. 74, p. 213 (1930).

|| Taylor, Meek and Jackson, *ibid.*, vol. 84, p. 373 (1933).

¶ Taylor, Darbyshire and Strunz, *ibid.* (*in press*).

\*\* Taylor and Jackson, *ibid.*, vol. 86, p. 53 (1933).



The feldspars show only the first kind of isomorphous replacement. The relative intensities of X-ray reflections show that the barium ions in celsian† occupy the same positions as the potassium ions in orthoclase, and that in hyalophanes of various compositions the barium ions replace potassium ions at random. These monoclinic feldspars thus form an isomorphous series of the simplest type. In the plagioclases we probably have an example of a double isomorphous series, one branch starting with the albite structure, the other with the anorthite structure. That the two structures differ to an appreciable extent is shown by the fact that in albite the *true* *c*-axis is 7.1 Å., whereas in anorthite this is only a pseudo-axis and the true axis is twice as long. Such evidence as is available\* suggests that the albite structure persists until one-half of the kations are calcium ions, and that the structure of crystals with more than this amount of calcium is that of anorthite, but this suggestion remains to be tested by further experimental work.

Zeolites provide examples of both types of isomorphous replacement. Hey and Bannister's analyses of optically-controlled material show that for thomsonite‡ it is necessary to assume that both replacements  $\text{NaSi} \rightleftharpoons \text{CaAl}$  and  $\text{Na}_2 \rightleftharpoons \text{Ca}$  have occurred. Since the sodium and calcium ions are nearly the same size it is natural that there should be isomorphous replacement of the first type, for it is always the same cavities which are occupied—either by sodium or by calcium ions. Replacement of the second type is possible because, of the eight cavities (per pseudo-cell) fitted to receive these kations, only six are occupied in normal thomsonite which contains  $\text{Na}_2\text{Ca}_4$ ; the maximum replacement possible is that which corresponds to a pseudo-cell with all its cavities occupied and therefore containing  $\text{Na}_6\text{Ca}_3$ . The material richest in sodium which was examined by Hey contained  $\text{Na}_{5.4}\text{Ca}_3$  approximately.

As we have seen in the previous section, the framework is essentially the same in natrolite, scolecite, and mesolite, and we may perhaps regard these three crystals as furnishing examples of isomorphous replacement of the type  $\text{Na}_2 \rightleftharpoons \text{Ca}$ . The unit cell of natrolite contains  $\text{Na}_{16}\text{Al}_{16}\text{Si}_{24}\text{O}_{80} \cdot 16\text{H}_2\text{O}$  and the 16 sodium ions occupy all the available cavities; the scolecite cell of the same size contains  $\text{Ca}_8\text{Al}_{16}\text{Si}_{24}\text{O}_{80} \cdot 24\text{H}_2\text{O}$  and only one-half of the cavities are occupied by calcium ions; and the mesolite cell is three times as large‡ and contains  $\text{Na}_{16}\text{Ca}_{16}\text{Al}_{48}\text{Si}_{72}\text{O}_{240} \cdot 64\text{H}_2\text{O}$  so that two-thirds of the cavities are occupied.

\* Taylor, Darbyshire and Strunz, *ibid.* (*in press*).

† 'Min. Mag.', vol. 23, p. 51 (1932).

‡ Hey and Bannister, 'Min. Mag.', vol. 23, p. 421 (1933).

Naturally-occurring crystals of the ultramarine type (sodalite, noselite, etc.) also furnish examples of isomorphous replacements of both types, but so far as I am aware the structure suggested by Pauling for sodalite does not indicate precisely how the second type of substitution proceeds.

In base-exchange experiments with zeolites, permutites, and ultramarines, ions occupying definite positions within the crystal lattice are removed from the lattice and others replace them, and with many zeolites the operation may be carried out without destroying the single crystal specimen. For this to be possible the tetrahedron framework of the crystal must be penetrated by channels or tunnels large enough to permit the movement of ions, and in the zeolites which have been analysed such tunnels have been found. In analcite they are parallel to the diagonals of the unit cube and are non-intersecting, in the ultramarines they are of the same size but intersect to form a large cavity within the framework, and in the fibrous zeolites they are parallel to the *c*-axis.

The water molecules which are included in the groups of oxygen-like atoms around kations in zeolites probably play a very important role in determining the possibility and relative facility of base-exchange. The water molecules must be easily moved (witness the dehydration and rehydration properties of zeolites) and for base-exchange to proceed it is probably necessary that the group around the kation shall include some water molecules. That the water molecules are concerned directly with the base-exchange is suggested very strongly by the nature of the group around the sodium ion in analcite. In this crystal the surrounding group contains four oxygen atoms, and two water molecules which lie in two non-intersecting channels through which ions and water molecules may move; the sodium ion itself lies midway between the two channels at their point of nearest approach.

X-ray measurements of certain substituted zeolites have been carried out. Reference has already (Section II) been made to the author's use of silver analcite\* in which the alterations in relative intensities of reflection show that the silver ions are actually occupying the places occupied by sodium ions in ordinary analcite. Similar calculations have been made for silver natrolite,† with the same result, and for lithium natrolite,† with results which suggest strongly but do not definitely prove that the same is also true in this case. Bannister's measurements‡ on substituted edingtonites prepared by Hey are included in Table I. In these substitutions the divalent barium ion is replaced

\* Taylor, 'Z. Kristallog.', vol. 74, p. 1 (1930).

† Taylor, Meek and Jackson, *loc. cit.*

‡ Hey and Bannister, private communication.

by two monovalent ions, a type of replacement possible because the barium ions in ordinary edingtonite occupy only one-half of the available cavities, while the monovalent ions in the substitution products probably occupy all these cavities. When silver replaces barium in edingtonite the crystalline needles give extremely weak X-ray pictures indicating that much amorphous material is present, and the sodium derivative gives an entirely different X-ray picture which shows that it is not a true edingtonite at all. The breakdown of the structure in these two products is probably a consequence of the replacement of large ions by much smaller ones.

Measurements on the substitution products of the platy zeolites heulandite and stilbite, in which tetrahedron sheets are probably held together largely by kation bonds, will be of great interest; and measurements on substituted chabazites in which the cavities are as large as those in the ultramarines and in which the base-exchange can be effected with great rapidity, may shed further light on the part played by the water molecule in the reaction.

#### IV. THE WATER IN ZEOLITES.

The peculiar dehydration and rehydration properties of zeolites, and the facility with which other molecules such as alcohols, mercury, and ammonia, can take the place of the water, have always been assumed to indicate that zeolitic water is structurally different from ordinary water of crystallization. The X-ray examination of zeolite structures described in this paper shows that the stability of the tetrahedron framework (and therefore of the whole crystal) is not dependent upon the water molecules, whereas the water molecules in salts such as  $\text{BeSO}_4 \cdot 4\text{H}_2\text{O}^*$  and  $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}^\dagger$  are just as important as the sulphate groups in determining the nature of the structure. The platy-zeolites, heulandite, and stilbite are probably exceptional in their dehydration properties, for approximately three-fifths of the normal water-content is removed with ease while the remainder appears to be more firmly held: this latter amount may perhaps represent water molecules which are necessary to the stability of the sheet structure in accordance with the suggestion advanced in Section II.

In the following paragraphs we shall discuss the nature of the forces between water molecules and adjacent atoms, the firmness of binding, and the replacement of water by other molecules.

\* Beevers and Lipson, 'Z. Kristallog.', vol. 82, p. 297 (1932).

*Ibid.*, vol. 83, p. 123 (1932).

The water molecules are always closely associated with the kations in zeolites, and always exhibit definitely polar properties, *e.g.*, the water molecule is never midway between two kations, but is attached either to one kation or else to two on the same side of the molecule. On examination of the relative positions of the nearest kations and oxygen ions it also appears that the binding forces must be very similar to those operating in ionic solutions in liquid water and in crystal structures such as beryllium sulphate and nickel sulphate mentioned above.

According to the interesting theory recently advanced by Bernal and Fowler\* the net electronic density distribution in the water molecule resembles a tetrahedron with two corners of positive and two of negative charge, *i.e.*, the water molecule may be supposed to possess four valency bonds, two positive and two negative, directed approximately towards the four corners of a regular tetrahedron. They also say that in an ionic solution the charge of an anion must be expected to attract the two hydrogen nuclei (positive bonds), that of a kation the two vacant places (negative bonds), symmetrically, and from the point of view of further co-ordination occupy them both.

Fig. 4 shows that in the zeolites so far examined the positive bonds of the water molecule are satisfied by oxygen ions, the negative bonds by either one divalent kation or two monovalent kations. Thus in natrolite the two negative bonds are satisfied by two sodium ions at 2.5 Å., the two positive bonds by two oxygen ions at 2.8 Å., and the four bonds are directed towards the four corners of a rather flattened tetrahedron. In edingtonite each water molecule of type  $Aq_1$  touches two oxygen ions at 2.8 Å., and one barium ion at 2.8 Å., each molecule  $Aq_2$  touches two oxygen ions at 3.1 Å. and one barium ion at 2.8 Å., and in both the bonds may be directed tetrahedrally if the kation is supposed to attract the two negative bonds symmetrically. In scolecite the arrangement of bonds is very similar, both negative bonds being satisfied by the calcium ion, and in thomsonite the molecules  $Aq_1$  are bound to two kations (probably one calcium and one sodium) and two oxygen ions with bonds arranged tetrahedrally as in natrolite. In all these systems an oxygen ion which satisfies one of the positive bonds of the water molecule does not touch a kation (it is shared by two tetrahedral  $SiO_4$  or  $AlO_4$  groups), and one may assume that any oxygen ion which belongs to two tetrahedra and also touches one or two kations has its valencies fully satisfied and cannot satisfy a positive bond from a water molecule. Adopting this assumption, the water molecule  $Aq_2$  in thomsonite is satisfied by one kation  $M_2$  (probably calcium) and two oxygen

\* 'J. Phys. Chem.,' vol. 1, p. 515 (1933).

ions which do not touch kations (while a third equally near oxygen ion is already fully satisfied), and the molecule  $Aq_3$  in the same crystal is similarly linked to the kation  $M_1$  (sodium or calcium) and two unsatisfied oxygen ions, two other oxygen ions at the same distance being already fully satisfied. Similarly in analcite, if the symmetry of the unit cell is supposed to be tetragonal, each water molecule touches two sodium ions and two unsatisfied water molecules in such a way as to satisfy its bonds supposed arranged tetrahedrally

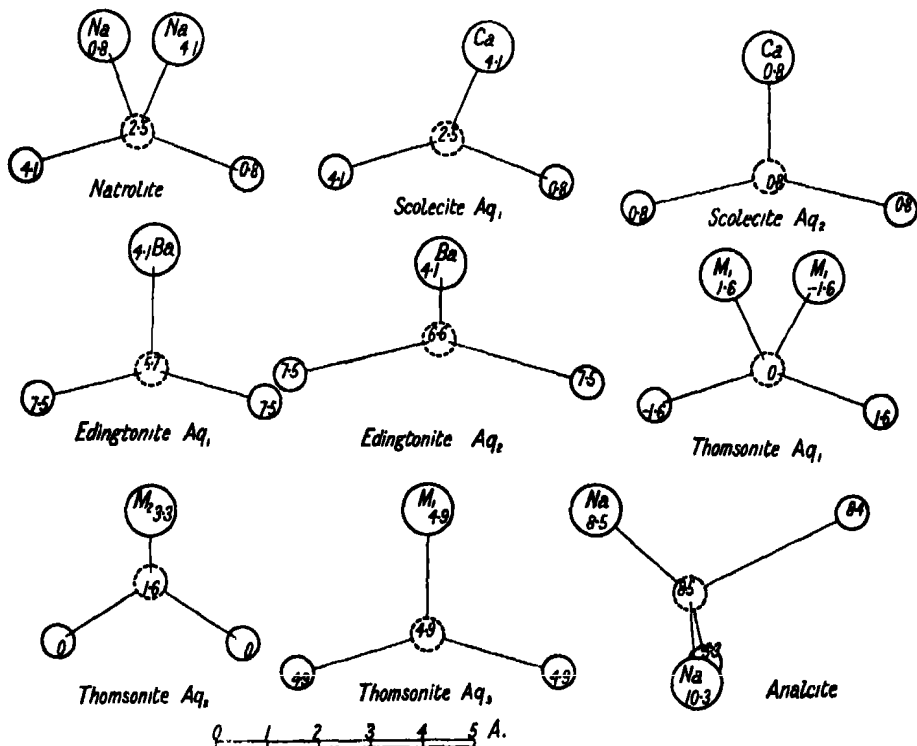


FIG. 4.—The binding of water molecules in zeolites. Water molecules are represented by dotted circles, oxygen atoms by small circles, and kations by large circles, and the heights of the atoms are given in A.

(see also Section V). The model of the water molecule proposed by Bernal and Fowler can therefore be applied to zeolitic water molecules in the structures which have been completely analysed, and it appears that the position taken up by the water molecule in a zeolite is such that its tetrahedrally directed positive and negative bonds are satisfied by negative and positive ions in the structure.

This hypothesis immediately suggests a simple explanation of Hey's observation that ammonium natrolite is the only anhydrous substituted natrolite

(see Table I); ordinary natrolite, if dehydrated, rapidly recovers its water on exposure to moist air, and other substituted natrolites are fully hydrated. If the ammonium ion ( $\text{NH}_4$ ) possesses four positive bonds directed tetrahedrally, these will be satisfied by the four nearest oxygen ions which form a flattened tetrahedral group around it; the kation cannot then satisfy the negative bonds of any water molecule and so the compound is anhydrous merely because there is no available ionic bond to attract the polar water molecule.

One other aspect of the binding of water molecules is worthy of remark, although its real significance is doubtful. In all the fibrous zeolites oxygen atoms which link adjacent chains form with water molecules double columns parallel to the *c*-axis of the crystal with a distance 3.3 Å. from oxygen atom to water molecule in the same column. Our incomplete investigation of heulandite and stilbite also suggests that water molecules may occupy positions between the tetrahedron sheets such that each molecule is midway between two oxygen atoms (one from each tetrahedron sheet) which are 6.6 Å. apart. The special significance of this distance 3.3 Å. is not clear, and it may in fact be an accidental consequence of the formation of co-ordination groups around kations and so possess no special significance at all.

The firmness of binding of water molecules in zeolites is measured by the molecular heat of hydration, large values indicating firm binding. The heat of hydration has been measured for thomsonite\* and natrolite† by Hey. In natrolite as the water-content falls from 16 (fully hydrated material) to 14 molecules per unit cell, the heat of hydration rises rapidly to about 25,000 cal. per gm. mol. and then remains nearly constant until dehydration is complete. In thomsonite the curve of heat of hydration against water-content shows three distinct regions corresponding to water-contents 12 to 8, 8 to 4, less than 4, water molecules per pseudo cell; for material with 4 molecules per pseudo cell the heat of hydration is about 30,000 cal. per gm. mol. It is reasonable to assume that the simple curve for natrolite is a consequence of the structural equivalence of all the water molecules, and that the three regions of the thomsonite curve are connected with the fact that in the structure there are three groups of four equivalent water molecules each; beyond this it does not appear possible to proceed until further measurements are available for known structures.

It is even more difficult to discuss in terms of structure the many interesting experiments made on the absorption of various gases by dehydrated zeolites.

\* Hey and Bannister, 'Min. Mag.', vol. 23 p. 51 (1932).

† *Ibid.*, p. 243.

Chabazite has probably been used for these experiments more frequently than any other zeolite, and we do not yet know the details of the chabazite structure\* ; but Hey has made some interesting observations on the absorption of alcohols by partially dehydrated thomsonite.† He suggests that alcohol molecules of length approximately 6-6.5 Å. are absorbed because the (OH) groups can satisfy residual affinities occurring at intervals equal to the length of the pseudo-axis  $c = 6.6$  Å., while longer alcohol molecules cannot do so ; the alcohol molecules are supposed to be arranged parallel to the  $c$ -axis of the crystal. Further experiments on carefully controlled material of known structure (particularly if the structure is less complex than that of thomsonite) will probably shed some light on the exact nature of the binding of the added molecules, but, granted channels sufficiently large to allow the passage of molecules through the framework, it seems probable that the absorption or non-absorption of a given molecule must be determined by considerations of bonding similar to those which fix the positions occupied by water molecules in the normal fully hydrated structure.

#### V. PSEUDOSYMMETRY AND TWINNING IN FRAMEWORK STRUCTURES.

Among the crystals with framework structures are many which are of particular interest because of special twinning properties or because of the close approximation of the external form of the crystal to that which would correspond to a higher class of symmetry than is actually possessed by the crystal. These properties reflect the geometrical properties of the tetrahedron framework, and must be explicable when the exact nature of the framework has been determined.

In the feldspars twinning is very frequently observed. The twins may be united in a great variety of ways, but it is possible to distinguish two main groups of twin laws which are consequences of different properties of the tetrahedron framework. In the first, repeated twinning of the triclinic plagioclases according to the Albite and Pericline laws produce pseudo-monoclinic complex crystals. Twinning of this type occur because the symmetry of the framework of the crystal, although strictly triclinic, departs only slightly from monoclinic ; a slight disturbance during the growth of the crystal may therefore result in the appearance, at one place in the complex crystal, of a symmetry element (reflection plane or rotation axis) characteristic of the very similar truly monoclinic related structure. The repeated twinning

\* See footnote § p. 82.

† Hey and Bannister, 'Min. Mag.', vol. 23, p. 243 (1932).

represents, as it were, the crystal's attempt to achieve a higher symmetry. Twinning of the second type (according to the Manebach, Carlsbad, and Baveno laws) are observed in both monoclinic and triclinic feldspars, and are a consequence of the special geometrical properties of the framework (essentially the same in all the feldspars). The feldspar framework has been described in Section II, and the tetrahedron strings from which it is constructed are illustrated in fig. 2. The three twin laws (Manebach, Carlsbad, Baveno) correspond to three different ways of securing that the tetrahedron strings at the place of union of the two individuals are practically unaltered in form while being linked so as to produce a continuous framework throughout the complex crystal.\* The frequency with which twinning takes place is presumably due to the fact that a very small disturbance is sufficient to force the atoms out of the arrangement which results in single-crystal growth, and into the arrangement which results in twin-growth.

The true symmetry of the orthoclase feldspars has always been doubtful. The plagioclases are dimensionally, optically, and structurally triclinic; many varieties of orthoclase are dimensionally and optically monoclinic, but microcline is both dimensionally and optically triclinic; and hyalophane is dimensionally and optically monoclinic while celsian is probably both dimensionally and optically triclinic. An examination of sanidine, hyalophane, and celsian has shown that structures of monoclinic symmetry explain very satisfactorily the observed relative intensities of X-ray reflection from these crystals. This must mean that if the symmetry of the individual unit cell is not accurately monoclinic, the departure from that symmetry is very slight by comparison, for example, with the departure in albite. Optical evidence that the symmetry is triclinic rather than monoclinic does not necessarily indicate a *large* departure of the atomic arrangement from the higher symmetry, because the optical method of examination is extremely sensitive by comparison with the X-ray method. It seems therefore that orthoclase and hyalophane are truly monoclinic, and that celsian is structurally monoclinic but is dimensionally and optically only pseudomonoclinic and actually triclinic (*cf.* natrolite below).

The fibrous zeolites exhibit a feature common in framework structures, viz., the tendency of the framework to have a high structural symmetry but a lower dimensional symmetry. Thus edingtonite is structurally tetragonal to a very close approximation† and so the X-ray intensities of reflection reveal this symmetry: but specimens from Bohlet are optically and dimensionally ortho-

\* Taylor, Darbyshire and Strunz (*loc. cit.*).

† Taylor and Jackson, 'Z. Kristallog.', vol. 86, p. 53 (1933).



rhombic ( $a = 9.56$  A.,  $b = 9.68$  A.,  $c = 6.53$  A.), while specimens from Old Kilpatrick are optically and dimensionally tetragonal ( $a = b = 9.59$  A.,  $c = 6.53$  A.).\* Similarly natrolite† is structurally tetragonal to a very close approximation, but is dimensionally and optically orthorhombic, and may be strictly monoclinic (certain very weak X-ray reflections are observed which indicate the lower symmetry). In scolecite the structural departure from tetragonal symmetry is still very slight, but optically and dimensionally the crystal is definitely monoclinic.

Thomsonite is an example of the much rarer converse case; here the structure is quite definitely orthorhombic with no approach to tetragonal symmetry at all, but dimensionally the crystal is tetragonal, for the axes  $a$  and  $b$  are the same length to within 0.1%.‡

Twinning is frequently observed in scolecite and is almost universal in mesolite. The frameworks of the crystals are almost tetragonal, and the tendency to twin growth represents the tendency to approximate to a symmetry higher than that of the individual unit cell (*cf.* feldspar twins, above).

Analcite§ is particularly interesting from the point of view of symmetry. The structure has been described in Section II, and we saw there that the X-ray reflections indicate that the symmetry is cubic. Optically analcite is often tetragonal or orthorhombic, a single (cubic) crystal being divided into sectors when examined under the microscope. I have examined such sectors cut from a single crystal, and find that all the sectors give identical X-ray rotation diagrams which show cubic symmetry; therefore, if the apparently cubic crystal is really a complex twin, the scale of the twinning must be much finer than is indicated by the optical division into sectors.

To arrive at a structure based on a unit cell with cubic symmetry it is necessary to assume first that all tetrahedra are identical, second that 16 sodium ions occupy uniformly the 24 available cavities. There is no serious difficulty in accepting the assumption that  $\text{AlO}_4$  groups and  $\text{SiO}_4$  groups are indistinguishable, but the second assumption must mean that the distribution of 16 ions in 24 holes is either a space average or a time average. If it is a space average, certain holes are always empty; if a time average, every cavity is occupied by an ion at some time, and it must be possible for a sodium ion to move from one cavity to another.

\* Hey and Bannister, private communication.

† Taylor, Meek and Jackson, 'Z. Kristallog.,' vol. 84, p. 373 (1933).

‡ Hey and Bannister, 'Min. Mag.,' vol. 23, p. 51 (1932).

§ Taylor, 'Z. Kristallog.,' vol. 74, p. 1 (1930).

X-ray methods cannot determine whether it is a space average or a time average, but the theory of the binding of water molecules discussed in Section IV does suggest why an average distribution is found. The framework is cubic, and the cavities for sodium ions are arranged in groups of three around the trigonal axis on which the water molecule lies. If *all* the cavities were occupied by sodium ions, each water molecule would touch three sodium ions and every oxygen atom throughout the structure would be linked to two silicon or aluminium ions and one sodium ion, and so could not satisfy the positive bond of a water molecule (in accordance with the hypothesis advanced in Section IV). If *two* of each group of three cavities are occupied by sodium ions, each water molecule touches *two* sodium ions which satisfy its negative bonds, and *two* oxygen ions which are now linked to two silicon or aluminium atoms only, and are therefore able to satisfy the positive bonds of the water molecule. The distribution of the sodium ions so that all the available cavities are occupied on the average may be ascribed to the fact that the tetrahedral bonds of the water molecule are satisfied by two sodium ions and two (unsatisfied) oxygen ions, when *any two* of the three cavities forming the group about the trigonal axis are occupied. The bond-properties of the water molecule demand that two of the cavities shall be occupied by sodium ions, and the third empty; the total number of sodium ions is thus fixed at 16 per unit cell and to balance valencies the number of aluminium ions must be 16 per unit cell also.

While leaving unsettled the question whether the average is in space or time, the above hypothesis does seem to account satisfactorily for the constant composition of analcite and for its pseudosymmetry.

It is legitimate, in a section devoted to pseudosymmetry and twinning, to refer to the parallel growths and intergrowths so frequently observed in crystals with framework structures. The parallel growth of the feldspars albite and orthoclase produces perthites and microperthites (which are lamellar crystals of widely-varying degrees of fineness of twinning) and is probably responsible for the "schiller" of moonstones. Among the fibrous zeolites intergrowths are common; parallel growths of natrolite with thomsonite, scolecite, and mesolite have been described by Hey,\* and intergrowths of thomsonite with mesolite and with edingtonite are also of fairly frequent occurrence. These forms of association are clearly due to the fact that all the crystals are built up from similar tetrahedron strings.

\* Hey and Bannister, 'Min. Mag.,' vol. 23, p. 243 (1932).

## VI. CONCLUSION.

In this paper attention is directed to some general properties which detailed analysis has shown to be characteristic of structures based on frameworks of linked tetrahedral  $\text{SiO}_4$  and  $\text{AlO}_4$  groups.

The fundamental unit is the tetrahedron of oxygen atoms, but all frameworks yet described contain as a secondary unit the ring of four tetrahedra, in which it is usually impossible to distinguish the very similar  $\text{SiO}_4$  and  $\text{AlO}_4$  groups. Three-dimensionally infinite frameworks of various types have been described, but it is suggested that the frameworks of the zeolites heulandite and stilbite may be of finite thickness, and infinite in only two dimensions.

Two types of isomorphous replacement of kations are observed in framework structures; in one the number of kations remains unaltered, in the other one divalent kation is replaced by (or replaces) two monovalent kations. Replacement of the second type is possible only when there are within the framework of the structure some cavities which may be either occupied or empty according as the total number of kations is large or small. Base exchange experiments with zeolites are also discussed.

The binding of the water molecules in zeolite structures is precisely such as will satisfy the tetrahedral positive and negative bonds which Bernal and Fowler have ascribed to the molecule in liquid water.

Crystals with framework structures show a marked tendency to form twin growths with symmetry higher than that of the individual unit cell, a tendency due in part to the fact that the form of the framework often approximates closely to that which would correspond to the higher symmetry. It is suggested that the peculiar pseudosymmetrical properties of the zeolite analcite are intimately connected with the nature of the binding of the water molecules within the framework.

I wish to express my gratitude to Professor W. L. Bragg, F.R.S., for his continued interest in the researches on which this paper is based, and for his helpful advice and constructive criticism at all stages of the work. I am also indebted to collaborators whose names appear in the published accounts of the analysis of individual structures, and to Messrs. M. H. Hey and F. A. Bannister for the very valuable help which they have so freely given throughout.

## SUMMARY.

The structures of feldspars and zeolites are based on strong frameworks of linked tetrahedral ( $\text{SiO}_4$ ) and ( $\text{AlO}_4$ ) groups, with kations and water molecules

in the interstices. In structures hitherto described the framework forms a three-dimensionally infinite anion, but in the platy zeolites heulandite and stilbite it is probable (though not certain) that the framework consists of separate sheets each of finite thickness.

In isomorphous replacement in the feldspars divalent kations replace monovalent kations of the *same* size, the same cavities are always occupied, and valencies are balanced by alteration of the proportion of aluminium to silicon, the replacements in orthoclase and plagioclase groups being represented by  $\text{KSi} \rightleftharpoons \text{BaAl}$  and  $\text{NaSi} \rightleftharpoons \text{CaAl}$  respectively. Similar replacements are found in zeolites, and in addition replacements such as  $\text{Ca} \rightarrow \text{Na}_2$ , a replacement possible when there are in the framework empty cavities which are occupied when sodium replaces calcium. In base-exchange experiments with zeolites both types of replacement can be produced without destroying the crystal structure because the frameworks are penetrated by channels large enough to permit the movement of kations and water molecules.

The peculiar dehydration properties of zeolites are due to these channels, and not to any special type of bonding, for the water molecules in zeolites take up such positions that their positive and negative bonds, arranged tetrahedrally, are satisfied by oxygen ions and kations respectively in much the same way as in ordinary hydrated salts.

In both feldspars and zeolites twinning and intergrowths are frequent because the frameworks are essentially simple in structure and often possess a high *structural* symmetry to which the actual *dimensional* symmetry approximates.

---

## *Stress Calculation for Tubular Frameworks having Continuous Longitudinals.*

By J. B. B. OWEN, B.Sc.

(Communicated by R. V. Southwell, F.R.S.—Received December 21, 1933.)

### *Introduction.*

1. In a recent paper Southwell\* has shown that the calculation of primary stresses for a tubular framework, generally representative of a rigid airship, is a problem which can be solved exactly by a synthesis of known "type solutions," although the order of redundancy is such as to make it quite intractable by conventional methods. By "primary stresses" is meant those actions which would be induced in the constituent members by loads applied solely at the joints, if the joints were entirely free, so that no bending actions could come into play. It is customary to assume such freedom in stress calculation,† both because the complexity of the problem is thereby reduced, and because primary stresses are in general predominant. But in proceeding to the design of the constituent members it is necessary to consider (at least in approximate fashion) what actions will be imposed as a result of rigidity in the joints; and in regard to the stressing of rigid airships it may be said that the outstanding problem (now that a technique exists for primary stress calculation) is the effect of continuity of the longitudinals.

The purpose of this paper is to show that Southwell's method can be extended so as to throw light on this question. Given the idea of a synthesis of known "type solutions," we have only to examine the nature of these type solutions when continuous longitudinals are presumed. We shall find that their nature is unchanged, the only difference being that the ratios of the component displacements, and the rate of their variation from bulkhead to bulkhead, are slightly altered; and since it is known from experience that a close approximation can be obtained by neglecting joint rigidity, we may solve the amended equations by treating the corrections as small quantities whose squares and products can be neglected.

In the nature of the case, rather lengthy numerical calculations will be required in any particular instance. These are left for subsequent discussion :

\* 'Proc. Roy. Soc.,' A, vol. 139, p. 475 (1933).

† *Ibid.*, § 2.

the present paper is intended merely to indicate the possibility of extending Southwell's method to a problem which is generally regarded as having importance, and acquaintance with his notation and results will be presumed.

### *Application of the Theorem of Three Moments.*

2. Let A, B, C be three consecutive points of support of a long straight girder having uniform flexural rigidity  $B$  throughout its length. Let  $\delta_A$ ,  $\delta_B$ ,  $\delta_C$  be the displacements of A, B, C below some datum level, measured *downwards*; and let  $M_A$ ,  $M_B$ ,  $M_C$  be the bending moments which come into existence at these points, taken positive when "hogging." When no forces are imposed on the girder except at the joints, Clapeyron's "Theorem of Three Moments" \* shows that

$$l_{AB} (M_A + 2M_B) + l_{BC} (2M_B + M_C) + 6B \left( \frac{\delta_B - \delta_A}{l_{AB}} + \frac{\delta_B - \delta_C}{l_{BC}} \right) = 0. \quad (1)$$

Suppose now that fig. 1 relates to the deflection of a longitudinal girder in a radial plane. We have

$$l_{AB} = l_{BC} = l \text{ (the length of a bay),}$$

so that (1) takes the form

$$M_A + 4M_B + M_C = -6 \frac{B_r}{l^2} (2\delta_B - \delta_A - \delta_C), \quad (2)$$

where  $B_r$  is the relevant flexural rigidity.

Also the shears in AB and BC, tending to move the supports at A and C upwards in relation to the support at B, are given by

$$V_{AB} = \frac{1}{l} (M_B - M_A)$$

and

$$V_{BC} = \frac{1}{l} (M_B - M_C),$$

so that the upward force which must be exerted on the girder by the support at B is

$$R_B = V_{AB} + V_{BC} = \frac{1}{l} (2M_B - M_A - M_C). \quad (3)$$

We now express these results in the notation of Southwell's paper. Taking the outward radial direction as "upward," we replace  $\delta_A$ ,  $\delta_B$ ,  $\delta_C$  by  $-u_A$ ,

\* Cf. (e.g.) J. Case, "Strength of Materials," § 194, equation (2), in which bending moments are taken as positive when "sagging."

$-u_B, -u_C$ ; and in conformity with his § 13, if  $x$  is the value assumed by some quantity at A, we write

$$\begin{aligned} & x \text{ for } x_A \\ & x + \Delta x \text{ for } x_B \\ & x + \Delta x + \Delta(x + \Delta x) \text{ for } x_C, \end{aligned} \quad (4)$$

so that

$$2x_B - x_A - x_C = -\Delta^2 x$$

Then (2) and (3) can be written in the form

$$\left. \begin{aligned} [\Delta^2 + 6(\Delta + 1)]M &= -6 \frac{B_r}{l^2} \Delta^2 u \\ [\Delta + 1]R &= -\frac{1}{l} \Delta^2 M \end{aligned} \right\} \quad (5)$$

3. In the standard type-solution which is investigated in Southwell's paper (§ 15) the radial and tangential components of displacement vary with the "bulkhead number"  $\beta$  in accordance with the law

$$u, v \propto \cosh(\mu\beta + \eta), \quad (6)$$

where  $\eta$  is arbitrary and  $\mu$  a constant which is to be determined. If we assume that  $M$  and  $R$  depend on  $\beta$  in the same manner, and use the relation (Southwell's equation (21))

$$[\Delta^2 - 2(\cosh \mu - 1)(\Delta + 1)] \cosh(\mu\beta + \eta) = 0, \quad (7)$$

we can throw the relations (5) into the equivalent form

$$\begin{aligned} (\cosh \mu + 2)[\Delta + 1]M &= -6 \frac{B_r}{l^2} (\cosh \mu - 1)[\Delta + 1]u, \\ [\Delta + 1]R &= -\frac{2}{l} (\cosh \mu - 1)[\Delta + 1]M, \end{aligned}$$

and then we have, in the notation explained by (4),

$$\begin{aligned} (\cosh \mu + 2)M_B &= -6 \frac{B_r}{l^2} (\cosh \mu - 1)u_B \\ R_B &= -\frac{2}{l} (\cosh \mu - 1)M_B. \end{aligned}$$

Hence (eliminating  $M_B$ ) we obtain

$$(\cosh \mu + 2)R_B = 12 \frac{B_r}{l^2} (\cosh \mu - 1)^2 u_B.$$

This is a relation between the radial displacement at B and the radial force which (for a solution of standard type) is required at B to maintain this dis-

placement against the resistance opposed by the flexural rigidity of the longitudinal. Since a similar relation must hold in respect of every joint, we may write (in general accordance with Southwell's notation)

$$(\cosh \mu + 2) \bar{R}_j = 12 \frac{B_r}{l^3} (\cosh \mu - 1)^2 u_j. \quad (8)$$

An exactly similar investigation shows that

$$(\cosh \mu + 2) \bar{T}_j = 12 \frac{B_t}{l^3} (\cosh \mu - 1)^2 v_j, \quad (9)$$

where  $\bar{R}_j$ ,  $\bar{T}_j$ , are the radial and tangential forces *required*, on account of the stiffness of the longitudinals, at the joint numbered  $j$  of any bulkhead. In (9),  $v_j$  is the tangential component of displacement, and  $B_t$  the appropriate value of the flexural rigidity, which in general will be different from  $B_r$ . We have assumed in the foregoing discussion that the principal axes of flexure for a longitudinal member lie in radial and tangential planes.

4. Equations (22) of Southwell's paper are the conditions for equilibrium in a mode specified by displacements

$$\left. \begin{aligned} u_j &= U_n \cos (nj\alpha + \epsilon) \cosh (\mu\beta + \eta) \\ v_j &= V_n \sin (nj\alpha + \epsilon) \cosh (\mu\beta + \eta) \\ w_j &= W_n \cos (nj\alpha + \epsilon) \sinh (\mu\beta + \eta) \end{aligned} \right\}. \quad (10)$$

When modified to allow for the effects of continuity of the longitudinals, as represented by (8) and (9), they become

$$\begin{aligned} & U_n \sin \frac{\alpha}{2} \left\{ k_4 \frac{(\cosh \mu - 1)^2}{\cosh \mu + 2} + k_3 (1 + \cos n\alpha) + 1 + \cos n\alpha \cosh \mu \right\} \\ & \quad + V_n \cos \frac{\alpha}{2} \sin n\alpha (k_2 + \cosh \mu) + \frac{l}{l} W_n \cos n\alpha \sinh \mu = 0 \\ & U_n \sin \frac{\alpha}{2} \sin n\alpha (k_2 + \cosh \mu) \\ & \quad + V_n \cos \frac{\alpha}{2} \left\{ k_5 \frac{(\cosh \mu - 1)^2}{\cosh \mu + 2} + k_2 (1 - \cos n\alpha) + 1 - \cos n\alpha \cosh \mu \right\} \\ & \quad + \frac{l}{l} W_n \sin n\alpha \sinh \mu = 0 \\ & U_n \sin \frac{\alpha}{2} \cos n\alpha \sinh \mu + V_n \cos \frac{\alpha}{2} \sin n\alpha \sinh \mu \\ & \quad + \frac{l}{l} W_n \{ \cos n\alpha \cosh \mu - 1 + k_1 (\cosh \mu - 1) \} = 0 \end{aligned} \quad (11)$$



where

$$k_1 = \frac{\Omega_t}{2\Omega_d}, \quad k_2 = \frac{\Omega_t}{2\Omega_d}, \quad k_3 = \frac{1}{18} \frac{\Omega_r}{\Omega_d} \operatorname{cosec}^4 \frac{\alpha}{2},$$

(as in the original paper), and  $k_4, k_5$  are new parameters defined by

$$k_4 = \frac{3B_r}{t^{2/3} \Omega_d \sin^2 \frac{\alpha}{2}}, \quad k_5 = \frac{3B_t}{t^{2/3} \Omega_d \cos^2 \frac{\alpha}{2}} \quad (12)$$

As in the original paper (equation (19)) we are to adopt the convention that

$$k_3 = 0, \text{ when } n = 1$$

5. It follows that Southwell's determinantal equation (26) is replaced by

$$\begin{array}{ccc} k_4 \frac{(s-1)^2}{s+2} + a, & h, & f \\ h, & k_5 \frac{(s-1)^2}{s+2} + b, & g \\ f, & g, & c \end{array} = D' = 0, \quad (13)$$

where, in conformity with his equations (26) and (50),

$$\begin{aligned} a &= 1 + k_2(1 + \cos n\alpha) + k_3 + s \cos n\alpha \\ b &= 1 + k_2(1 - \cos n\alpha) - s \cos n\alpha \\ c &= -1 + k_1(s-1) + s \cos n\alpha \\ f &= \cos n\alpha \sqrt{(s^2-1)} \\ g &= \sin n\alpha \sqrt{(s^2-1)} \\ h &= \sin n\alpha (k_2 + s) \end{aligned} \quad \left. \vphantom{\begin{aligned} a \\ b \\ c \\ f \\ g \\ h \end{aligned}} \right\}. \quad (14)$$

and  $s$  is written for  $\cosh \mu$

Expanding the determinant, we have

$$D' = \frac{(s-1)^4}{(s+2)^2} ck_4k_5 + \frac{(s-1)^2}{s+2} \{k_4(bc - g^2) + k_5(ac - f^2)\} + D = 0, \quad (15)$$

where  $D$  is the determinant

$$\begin{vmatrix} a & h & f \\ h & b & g \\ f & g & c \end{vmatrix}$$

which is expanded in Southwell's equation (27). That equation, which holds in relation to pin-jointed longitudinals, is replaced by (15) when the longitudinals are continuous.

6. Equation (15) is a quintic involving five variable "elasticity parameters"  $k_1, \dots, k_5$ , and in it  $n$  may have any integral value from 0 to  $\frac{1}{2}N$ . A general solution is out of the question\*; and while numerical solutions can be obtained for a particular framework (when values have been assigned to the  $k$ 's) the labour involved is hardly justified, on account of the relative smallness of  $k_4, k_5$ .

When the longitudinals have a uniform cross-section, it is easy to show that the flexural rigidities can be expressed in the forms

$$B_r = l^3 \Omega_r \kappa_R^2, \quad B_t = l^3 \Omega_t \kappa_T^2,$$

where  $\kappa_R, \kappa_T$  are the radii of gyration of the cross-section for bending in the radial and tangential plane. Hence we have from (12)

$$\left. \begin{aligned} k_4 &= \frac{3 \Omega_r \kappa_R^2}{\Omega_d l^2 \sin^2 \frac{\alpha}{2}} = k_1 \frac{6 \kappa_R^2}{l^2 \sin^2 \frac{\alpha}{2}} \\ k_5 &= \frac{3 \Omega_t \kappa_T^2}{\Omega_d l^2 \cos^2 \frac{\alpha}{2}} = k_1 \frac{6 \kappa_T^2}{l^2 \cos^2 \frac{\alpha}{2}} \end{aligned} \right\}. \quad (16)$$

Accordingly, in practice,  $k_4$  and  $k_5$  will be small in relation to  $k_1$ , and the first term in  $D'$ , which involves their product, can be neglected. In these circumstances equation (15) reduces to a *quartic*.

Theoretically, five values of  $s$  (leading to five distinct solutions for any one value of  $n$ ) are necessary for the satisfaction of the imposed conditions, just as three were necessary in Southwell's problem. In that problem, at either end of the framework, three component forces (or displacements) could be specified; here, in addition, two component couples (or slopes) can be specified as terminal conditions for the longitudinal girders. But it is evidently not of great importance that we should satisfy these additional conditions exactly; what is of importance is to trace the effect of continuity in the longitudinals upon the general distortion of the tube.

7. We must, however, before passing to the simplified equation, examine the possibility that the exact equation (15) has a repeated root ( $s = 1$ ), because in this event the number of available solutions will *not* be sufficient. Evidently the condition is that ( $s = 1$ ) shall be a repeated root of Southwell's equation

$$D = 0,$$

\* In the absence of radial bracing ( $k_2 = 0$ ) the left-hand side of (15) will factorize into a cubic expression multiplied by  $(s - 1)^2$ .

and it was shown in his paper (§ 17) that this will be true when either  $\cos n\alpha = 1$  ( $n = 0$ ) or  $k_3 = 0$  (i.e., by (12), when  $n = 1$ ). So in relation to these cases a fresh examination of the difference equations is necessary.

8. If we assume that

$$\left. \begin{aligned} \bar{R}_j &= R_n \cos (nj\alpha + \varepsilon) \\ \bar{T}_j &= T_n \sin (nj\alpha + \varepsilon) \end{aligned} \right\}. \quad (17)$$

the first and second of Southwell's difference equations (18) are to be modified by the addition of the terms

$$\frac{1}{2l^2 \Omega_d \sin \frac{\alpha}{2}} [\Delta + 1] R_0 \quad \text{and} \quad \frac{1}{2l^2 \Omega_d \cos \frac{\alpha}{2}} [\Delta + 1] T_0,$$

respectively. Thus when  $n = 0$  we have

$$\begin{aligned} [\Delta + 1] R_0 \frac{1}{2l^2 \Omega_d \sin \frac{\alpha}{2}} + [\Delta^2 + 2(2 + 2k_2 + k_3)(\Delta + 1)] U_n \sin \frac{\alpha}{2} \\ + \frac{l}{t} [\Delta^2 + 2\Delta] W_0 = 0 \\ [\Delta + 1] T_0 \frac{1}{2l^2 \Omega_d \cos \frac{\alpha}{2}} - \Delta^2 V_0 \cos \frac{\alpha}{2} = 0 \end{aligned} \quad (18)$$

$$[\Delta^2 + 2\Delta] U_0 \sin \frac{\alpha}{2} + (1 + k_1) \frac{l}{t} \Delta^2 W_0 = 0$$

which may be compared with equations (31) of Southwell's paper. Also from (5), eliminating  $M$ , we have

$$[\Delta^2 + 6(\Delta + 1)][\Delta + 1] R_0 = 6 \frac{B_r}{l^3} \Delta^4 U_0 \quad \left. \vphantom{[\Delta^2 + 6(\Delta + 1)]}$$

and similarly

$$[\Delta^2 + 6(\Delta + 1)][\Delta + 1] T_0 = 6 \frac{B_t}{l^3} \Delta^4 V_0 \quad (19)$$

as equations replacing (8) and (9), which hold in relation to type solutions of the form (10).

9. The second of (18) and (19) are independent of the others. Eliminating  $T_0$ , we obtain

$$[(k_3 - 1) \Delta^2 - 6(\Delta + 1)] \Delta^2 V_0 = 0,$$

whence we have either

$$[(k_3 - 1) \Delta^2 - 6(\Delta + 1)] V_0 = 0,$$

leading to a solution of the normal type (10) in which

$$(k_s - 1)(s - 1) = 3,$$

or

$$\Delta^2 V_0 = 0,$$

with a solution

$$V_0 = P + Q\beta,$$

which (cf. § 21 of Southwell's paper) represents a rotation of the framework, combined with uniform twist. The longitudinals are not bent,\* and hence their flexural rigidity has no effect.

In the same way, eliminating  $R_0$  and  $W_0$  from the first of (19) and from the first and third of (18), we arrive at an equation which gives two solutions of standard type, together with a solution in which

$$\Delta U_0 = \Delta^2 W_0 = 0.$$

In this solution also the rigidity of the longitudinals plays no part: the displacements involve a rigid-body translation combined with uniform *stretch*.

Thus, in the case  $n = 0$ , no new consideration arises except in relation to solutions of the standard type.

10. When  $n = 1$ ,  $k_s$  is to be treated as zero, and  $(s - 1)$  again appears as a repeated root of (15). It can be shown that  $(s - 1)^3$  is not a factor of  $D'$ , and hence we have, as in the case ( $n = 0$ ), to find four, and only four, solutions of non-standard type.

Now in §§ 24-29 of Southwell's paper, when the corresponding case of his problem was discussed, the four solutions were found to be:—

- (i) a rigid body translation,
- (ii) a rigid body rotation,
- (iii) a flexural solution,
- (iv) a solution representing uniform shear.

The first two solutions involve no strain, and so can occur when the longitudinals are continuous. The third and fourth were shown to involve displacements given by

$$\begin{aligned} U_1 &= (1 - \cos \alpha)(1 + k_2)(F\beta^3 + G\beta^2) + 3F\beta + G \\ -V_1 &= (1 - \cos \alpha)(1 + k_2)(F\beta^3 + G\beta^2) \\ -\frac{l}{r}W_1 &= (1 - \cos \alpha)(1 + k_2)(3F\beta^2 + 2G\beta) \\ &\quad + 3F\{k_1(1 + k_2) + k_2 \cos \alpha\} \end{aligned} \tag{20}$$

\* To the order first of small quantities; cf. the first of equations (5).

11. Since

$$\begin{aligned}\Delta\beta &= 1, & \Delta\beta^2 &= 2\beta + 1, & \Delta\beta^3 &= 3\beta^2 + 3\beta + 1, \\ \Delta^2\beta &= 0, & \Delta^2\beta^2 &= 2, & \Delta^2\beta^3 &= 6(\beta + 1),\end{aligned}$$

we have according to (20)

$$\Delta^2 U_1 = -\Delta^2 V_1 = (1 - \cos \alpha) (1 + k_2) \{6(\beta + 1) F + 2G\},$$

and the first of (5) is satisfied if we write

$$u = U_1 \cos (j\alpha + \epsilon), \quad M = M_1 \cos (j\alpha + \epsilon),$$

where

$$M_1 = -\frac{B_r}{l^2} (1 - \cos \alpha) (1 + k_2) (6\beta F + 2G).$$

Then we see from the second of (5) that radial force is not required to maintain the displacements (20) in the longitudinals, and in the same way we can show that tangential force is not required. The longitudinals are subjected to uniform shears throughout their length, and these must be contributed by the terminal forces; but with this proviso it may be said that *none of the non-standard solutions described in § 10 are affected by continuity of the longitudinals.*

12. Reverting to the general equation (15), we now proceed (in accordance with the results of § 6) on the basis that  $k_4$  and  $k_5$  are small quantities (of the same order) in comparison with  $k_1$ . On this understanding three of the roots of (15) will differ by small quantities from the roots of Southwell's equation

$$D = 0,$$

which applies to pin-jointed longitudinals; and two new roots will make an appearance, corresponding with modes of distortion which involve bending of the longitudinals accompanied by relatively small displacements of the joints.

13. Considering the new roots first, we observe that the terms in (15) will be of the same order if  $(s + 2)$  is small of the order of  $k_4$  and  $k_5$ . Writing

$$s + 2 = \delta,$$

and neglecting small quantities of order  $k_4$ ,  $k_5$ ,  $\delta$ , we have from (15)

$$3^4 c \frac{k_4 k_5}{\delta^2} + 3^2 \left\{ \frac{k_4}{\delta} (bc - g^2) + \frac{k_5}{\delta} (ac - f^2) \right\} + D = 0, \quad (21)$$

where, in the expressions for  $a$ ,  $b$ ,  $c$ ,  $f$ ,  $g$ , and  $D$ ,  $s$  is to be given the value  $-2$ . The expansion of  $D$  (cf. equation (27) of Southwell's paper) is given by

$$\begin{aligned}-D &= (s - 1)^2 \{k_1 (s - 1) + 2k_1 (1 + k_2) + k_2 (1 + \cos n\alpha)\} \\ &\quad + k_2 [(1 + k_1 \cos n\alpha) (s - 1)^2 \\ &\quad \quad - \{k_1 (1 + k_2) - 2 + k_2 \cos n\alpha\} (1 - \cos n\alpha) (s - 1) \\ &\quad \quad \quad + (1 + k_2) (1 - \cos n\alpha)^2], \quad (22)\end{aligned}$$

and  $a, b, c, f, g$  are given in (14). So in (21) we must substitute

$$\left. \begin{aligned} a &= 1 + k_2 + k_3 + (k_2 - 2) \cos n\alpha \\ b &= 1 + k_2 - (k_2 - 2) \cos n\alpha \\ c &= (1 + 3k_1 + 2 \cos n\alpha) \\ f^2 &= 3 \cos^2 n\alpha \\ g^2 &= 3 \sin^2 n\alpha \\ -D &= 9 \{ -k_1 + 2k_1k_2 + k_2(1 + \cos n\alpha) \\ &\quad + k_3[9(1 + k_1 \cos n\alpha) \\ &\quad + 3\{k_1(1 + k_2) - 2 + k_2 \cos n\alpha\}(1 - \cos n\alpha) \\ &\quad + (1 + k_2)(1 - \cos n\alpha)^2] \} \end{aligned} \right\}, \quad (23)$$

and then the two values of  $\delta$  can be determined from the quadratic

$$D\delta^2 + 9\delta\{k_4(bc - g^2) + k_5(ac - f^2)\} + 81ck_4k_5 = 0. \quad (24)$$

14. The other three roots may be investigated when the roots of the cubic equation  $D = 0$  are known. Let  $s'$  stand for any one of these roots, and  $(s' + \delta')$  for the corresponding root of (15). When  $s$  has this latter value,  $D$  is given (to the first order of small quantities) by

$$\begin{aligned} & -[(s' - 1)\{3k_1(s' - 1) + 4k_1(1 + k_2) + 2k_2(1 + \cos n\alpha) \\ & \quad + k_3\{2(s' - 1)(1 + k_1 \cos n\alpha) \\ & \quad - [k_1(1 + k_2) - 2 + k_2 \cos n\alpha](1 - \cos n\alpha)\}\}] \delta', \end{aligned} \quad (25)$$

and the other terms in (15) become (to the same order, and hence neglecting the term involving  $k_4k_5$ )

$$\frac{(s' - 1)^2}{s' + 2} \{k_4(b'c' - g'^2) + k_5(a'c' - f'^2)\}, \quad (26)$$

where  $a', b', c', f', g'$  stand for  $a, b, c, f, g$  as given by (14) but with  $s'$  written for  $s$ .

Equating the sum of (25) and (26) to zero, we can calculate  $\delta'$  and hence determine the required root. We know in advance (from §§ 8 and 9) that attention need not be directed to the repeated root ( $s = 1$ ) which can occur when  $n = 0$  or  $1$ .

15. The possibility discussed in §§ 30-35 of Southwell's paper—of a repeated root other than  $s = 1$ —is so remote as hardly to call for notice in this paper. If in any numerical case it is found to occur, it may be dealt with on the lines which he has described.

A full discussion of our present problem is in fact not worth while, on account of the immense simplification that can be effected by inserting numerical values. What is significant in the present discussion is that all modes of distortion which imply *resultant* actions (namely, those associated with the root  $s = 1$  in the cases  $n = 0$  and  $n = 1$ ) have been shown to be unaltered by the continuity of the longitudinals, except in so far as additional forces are required to bend these; that the other modes (of standard type) are affected by small amounts which can be calculated; and that the new modes (if it is thought necessary to examine them) can be investigated without difficulty, on lines which have been explained in § 13.

#### *Summary.*

The effects of rigid joints in frameworks can be studied by processes involving continued approximations (*e.g.*, the method of Professor Hardy Cross), but do not appear to have been brought within the scope of direct calculation. The present paper is intended to indicate a direction in which some advance can be made. For the calculation of primary stresses (*i.e.*, the actions which would be induced if the joints were free) Southwell has recently suggested an exact method involving a synthesis of known "type-solutions," and it will sometimes be possible to take account of joint rigidity in relation to the latter. Where this can be done, an exact solution for specified loads, though laborious, will be practicable.

As an example, the effects of continuous longitudinals are considered in relation to a tubular framework of the kind discussed in Southwell's paper. (This case has importance in relation to the design of rigid airship hulls, where the effects of continuity of the longitudinals is generally regarded as the main outstanding problem.) It is shown that the more important of the type-solutions will be sensibly unaffected, and that the remainder will retain their general form. In practice the corrections will usually be small, and as such can be investigated without difficulty when numerical values have been assigned to the elastic parameters. On account of the number of these, detailed discussion of the general equations is hardly worth while.

---

*An Optical Examination of Thin Films. I.—The Optical Constants of Mercury.*

By LEIF TRONSTAD, Lecturer in Inorganic Chemistry, The Technical College, Trondheim, and C. G. P. FEACHEM, Denman Baynes Student, Clare College, Cambridge.

(The Laboratory of Colloid Science, Cambridge.)

(Communicated by E. K. Rideal, F.R.S.—Received December 22, 1933.)

*Introduction.*

Drude\* originally suggested that the properties of films on both absorbing and non-absorbing substrates could be studied by an optical method, since the reflecting properties of a surface are modified by the presence of a film. This method has in part been used by various workers,† with some success. For liquid surfaces, or for smooth surfaces of solids, such as are obtained by cleavage or after suitable polishing, it is possible to measure "the relative phase retardation,"  $\Delta$ , and "the ratio of the reflection coefficients,"  $\tan \psi$ , which define the reflecting properties.

From the values of  $\Delta$  and  $\psi$  for a clean surface, the refractive index,  $n$ , and the absorption coefficient,  $k$ , characteristic of the material can be found from the approximate equations

$$n = \sin \phi \tan \phi \cos 2\psi / (1 + \cos \Delta \sin 2\psi), \quad (1.1)$$

$$k = \sin \Delta \tan 2\psi, \quad (2.1)$$

where  $\phi$  is the angle of incidence.

The optical properties of a medium are often described by two quantities,  $\Phi$  and  $\Psi$ .  $\Phi$  is the principal angle of incidence at which angle  $\Delta = \pi/2$ , and  $\Psi$ , the principal azimuth, is the value of  $\psi$  at this angle of incidence. Solving equations (1.1) and (2.1),

$$2\Psi = \tan^{-1} k \quad (3.1)$$

$$\sin \Phi \tan \Phi = n / \cos 2\Psi. \quad (4.1)$$

Equation (4.1) may conveniently be solved by trial and error.

\* "The Theory of Optics," London, 1902, p. 363.

† See, for a complete bibliography, Tronstad, 'Trans. Faraday Soc.,' vol. 29, p. 502 (1933).



For a polished surface covered with a homogeneous film, if the optical properties of the clean surface be  $\bar{\Delta}$  and  $\bar{\psi}$ , and those of the surface covered by the film be  $\Delta$  and  $\psi$ , then if the film thickness be  $L$  and its refractive index  $n_1$ , Drude showed that to a first approximation, on an absorbing medium in air,

$$\Delta - \bar{\Delta} = -\frac{4\pi L}{\lambda} \frac{\cos \phi \sin^2 \phi}{(\cos^2 \phi - a)^2 + a'^2} \left(1 - \frac{1}{n_1^2}\right) (\cos^2 \phi - a) \quad (5.1)$$

$$2(\psi - \bar{\psi}) = \frac{4\pi L}{\lambda} \sin 2\bar{\psi} \frac{a' \cos \phi \sin^2 \phi}{(\cos^2 \phi - a)^2 + a'^2} \left(1 - n_1^2 \cos^2 \phi\right) \left(1 - \frac{1}{n_1^2}\right), \quad (6.1)$$

where

$$a = \frac{1 - k^2}{n^2(1 + \bar{k}^2)^2} \quad \text{and} \quad a' = \frac{2k}{n^2(1 + \bar{k}^2)^2},$$

and  $\lambda$  is the wave-length of light used.

On a non-absorbing medium, with a refractive index  $n$ , in air :

$$\tan \psi = \tan \bar{\psi} \text{ and is constant} \quad (7.1)$$

$$\Delta - \bar{\Delta} = \frac{4\pi L}{\lambda} \frac{\cos \phi \sin^2 \phi}{1 - \frac{1}{n^2} - \sin^2 \phi \left(1 - \frac{1}{n^4}\right)} (1 - n_1^2) \left(\frac{1}{n_1^2} - \frac{1}{n^2}\right). \quad (8.1)$$

Not only can the optical properties  $n_1$  and  $L$  of a film be determined from measurements of  $\Delta$  and  $\psi$ , but as Strachan\* has shown we may also evaluate the scattering indices of the film.

### *Experimental Method.*

The principle of the method adopted for measuring both  $\Delta$  and  $\psi$  consists essentially in reflecting an accurately parallel beam of elliptically polarized light of known properties from the surface at a known angle and in varying the ellipticity of the incident beam until the reflected light is plane-polarized. The azimuth of the reflected light is measured with a Lippich half-nicol, whilst the state of polarization is detected with a Szivessy half-shade system of special construction ; this divides the field of view into two parts which can only be matched in accurately plane-polarized light ; the details of the construction and use of this half-shade have been published by one of us (L. T.)† elsewhere, but the optical principles of the method there set out may be briefly summarized here. The plane of incidence is defined as lying in azimuth zero. A polarized nicol prism produces a beam of light, plane-polarized in azimuth  $P$  : this passes through a  $\frac{1}{4}$ -wave plate in azimuth  $\pi/4$ , and is converted

\* 'Proc. Camb. Phil. Soc.,' vol. 29, p. 116 (1933).

† 'J. Sci. Instr.,' in press (1934).

into a beam of elliptically polarized light of known azimuth and ellipticity. This is reflected from the surface and illuminates the ellipticity half-shade. The angle  $P$  can be varied by rotating the nicol until the light reflected from the surface is plane-polarized; when this is so, the two halves of the ellipticity half-shade appear equally bright. When this reflected beam is plane-polarized it can be shown\* that

$$\sin \Delta = \cos 2P, \quad (5.1)$$

and from this equation  $\Delta$  may be calculated. The ellipticity half-shade is then removed from the reflected beam, and the Lippich half-nicol put in its place, and the azimuth,  $A$ , of the reflected beam measured.  $\psi$  is then calculated from the equation

$$\tan \psi \tan A = 1. \quad (6.1)$$

The optical parts were mounted on a specially constructed Jamin Circle, consisting essentially of a polarizing spectrometer, with the particular feature that the arms moved in an accurately vertical plane to permit of the examination of liquid surfaces. The base consisted of a heavy tripod on levelling screws supported on three shock absorbers of felt, soaked in heavy oil, and mounted on a stone table independent of the floor in a cellar.

On one of two ribbed cast gun-metal arms was mounted a cadmium mercury lamp with glass filters providing monochromatic illumination of  $\lambda = 5780$ ,  $5460$ , and  $4358$  Å. respectively. The light from the lamp after passing through a collimator containing a diaphragm, lenses, a Glan-Thomson polarizing nicol ( $20 \times 11 \times 11$  mm.) and a Sénarmont compensator (quarter-wave mica plate), likewise mounted on the arm, falls on the reflecting surface. The azimuth of the nicol is read by means of a divided circle and vernier provided with a fine adjustment, whilst the angle of incidence of the light on the horizontal surface can likewise be measured with the aid of a vertical divided circle. The reflected light is received through a system mounted on the other gun-metal arm, comprising the Lippich half-nicol and the ellipticity half-shade system.

As will be mentioned later, the surfaces examined were not always homogeneous, so that the half-shades could not be used. To examine such surfaces, a microscope was built which could be focussed, through the analysing nicol, on to the surface itself. Different patches of surface appeared with different brightnesses, and by turning both nicols, the patches could each be extinguished

\* 'K, norske Vidensk. Selsk. Skr.,' No. 1, pp. 38, 207 (1931).

in turn. The values of  $P$  and  $A$  found by this method could be used in equations (5.1) and (6.1), and gave values of  $\Delta$  and  $\psi$  which, though less accurate than those obtained by the use of the half-shades, were useful as showing the relative properties of different parts of inhomogeneous surfaces.

In order to examine mercury surfaces, a chamber was constructed to enable the surface to be prepared and kept in clean dry nitrogen. The chamber was in the form of a horizontal cylinder of brass with two long slits round its curved face, covered with curved glass windows, ground to fit, and set in tap-grease. The chamber had outlets to connect with a source of nitrogen, a pump, and a still to produce mercury, and an opening covered with a rubber membrane through which the film material could be introduced with the aid of a hyperdermic needle.

These glass windows were, of course, liable to exhibit double refraction owing to internal strain, but would not be expected to have any effect on  $\psi$ , or on the angle of incidence, if they were set accurately normal to the beam. This accidental double refraction,  $\delta$ , might be in any azimuth,  $\theta$ , and would alter the apparent double refraction  $\Delta$ , of the surface, which is in azimuth zero.

The error introduced into the measured quantity  $\Delta$ , owing to this accidental double-refraction, would depend not only on  $\delta$  and  $\theta$  but also on the value of  $\Delta$ . Now it is not possible to measure both  $\delta$  and  $\theta$  without being able to rotate the windows, which was impracticable on this instrument, but it is easy to show that, if  $\delta$  be small, the error introduced into  $\Delta$  does not depend on  $\Delta$  within the limits of accuracy of the measurement, so that it can be determined for one value of  $\Delta$ , and applied as a correction to all values. This correction, the "window correction" was measured by observing the reflection from a stainless steel mirror with and without the cover; it was found to be about  $1^\circ$  and to remain fairly constant in time, and to return to the same value within 15 minutes of the chamber being evacuated and refilled. This last provision is, of course, essential.

The mercury was cleaned and distilled *in vacuo* directly from a pyrex still into a vitreosil dish supported at the appropriate level within the chamber. During the measurements the chamber was kept filled with nitrogen which had been carefully purified by passage over hot copper turnings, and then through a small furnace, containing sodium vapour, to remove the last traces of oxygen. It was then passed through sulphuric acid and olive oil, to remove dust (especially of sodium oxide), and water, and finally over phosphoric acid and through a further dust-filter of cotton-wool.

*Experimental Results.*

In all, twelve series of observations were made and  $\Delta$  and  $\psi$  measured (in centesimal degrees:  $1^\circ = \pi/200$ ) for various values of the angle of incidence,  $\phi$ , and for light of three wave-lengths. The results of three typical series, embracing all the angles of incidence used, are given in Tables I, II, and III, the values of  $\Delta$  being corrected for the window-effect. The observations in Table I were made at the beginning, immediately after the apparatus had been cleaned and assembled; those in Tables II and III during the course of the work, in the order in which they are given. Towards the end of the series, traces of contamination could be seen on the surface; these consisted of grey streaks, too thin to give interference colours, of some impurity which did not spread uniformly over the surface.

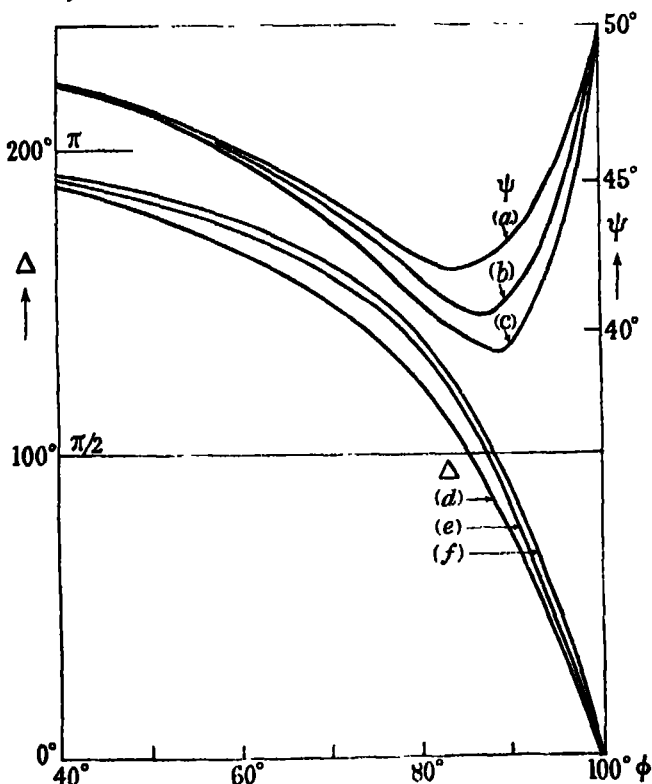


FIG. 1.— $\psi$  (a) Blue; (b) green; (c) yellow.  $\Delta$  (d) Blue; (e) green; (f) yellow.

No detectable contamination was present on the surface during the first series of observations, and the  $\Delta$ -values are higher than any subsequently observed and must be regarded as the best values obtained. They are plotted in fig. 1, and it will be seen that the form of the curves obtained is in general

Table I.

$\phi$ .	Yellow: $\lambda = 5780 \text{ \AA}$ .				Green: $\lambda = 5460 \text{ \AA}$ .				Blue: $\lambda = 4358 \text{ \AA}$ .			
	$\bar{\Delta}$ .	$\bar{\psi}$ .	$n$ .	$k$ .	$\bar{\Delta}$ .	$\bar{\psi}$ .	$n$ .	$k$ .	$\bar{\Delta}$ .	$\bar{\psi}$ .	$n$ .	$k$ .
$^{\circ}$	$^{\circ}$	$^{\circ}$			$^{\circ}$	$^{\circ}$			$^{\circ}$	$^{\circ}$		
95	50.07	42.81	1.68	3.08	48.08	44.02	1.38	3.61	40.4	45.3	1.04	3.99
90	90.18	39.78	1.67	2.97	86.67	40.87	1.41	3.31	74.9	42.9	1.01	4.16
80	138.00	41.01	1.75	2.85	134.85	41.66	1.52	3.19	120.8	42.2	1.03	3.79
70	161.83	43.68	1.79	2.81	159.55	44.06	1.55	3.14	149.2	44.3	1.00	3.95
60	175.54	45.65	1.85	2.73	174.30	45.94	1.61	3.06	167.2	46.0	1.04	3.90
50	184.80	47.06	1.90	2.61	183.99	47.23	1.75	2.83	179.7	47.0	(1.41)	(3.31)
40	190.79	48.05	(2.18)	(2.35)	190.72	48.13	(2.04)	(2.47)	189.1	48.1	(1.60)	(2.85)
Mean values of $n$ and $k$				2.84	—	—	1.61	3.10	—	—	1.03	3.95

Table II.

$\phi$ .	Yellow: $\lambda = 5780 \text{ \AA}$ .				Green: $\lambda = 5460 \text{ \AA}$ .				Blue: $\lambda = 4358 \text{ \AA}$ .			
	$\Delta$ .	$\bar{\psi}$	$n$ .	$k$ .	$\Delta$ .	$\bar{\psi}$ .	$n$	$k$	$\Delta$ .	$\bar{\psi}$ .	$n$ .	$k$ .
$^{\circ}$	$^{\circ}$	$^{\circ}$			$^{\circ}$	$^{\circ}$			$^{\circ}$	$^{\circ}$		
95	50.03	42.83	1.63	3.09	47.58	43.98	1.39	3.55	39.6	45.0	1.09	3.67
90	90.12	39.78	1.72	2.98	86.69	41.13	1.43	3.41	73.9	42.9	1.14	3.52
80	137.61	40.98	1.76	2.85	134.26	41.65	1.51	3.20	121.4	42.4	0.99	3.88
70	161.11	43.66	1.75	2.85	158.42	44.07	1.48	3.23	149.1	44.1	1.02	3.84
60	174.47	45.66	1.72	2.91	173.38	45.86	1.54	3.11	166.2	46.1	0.95	4.11
50	183.78	47.08	1.78	2.74	182.89	47.27	1.53	3.09	178.1	47.1	1.03	3.69
40	190.13	48.04	(1.92)	(2.51)	189.89	48.09	(1.78)	(2.63)	187.8	48.0	(1.31)	(3.03)
Mean values of $n$ and $k$			1.73	2.90	—	—	1.48	3.16	—	—	1.04	3.79

Table III.

$\phi$ .	Yellow: $\lambda = 5780 \text{ \AA}$ .				Green: $\lambda = 5460 \text{ \AA}$ .				Blue: $\lambda = 4358 \text{ \AA}$ .			
	$\bar{A}$ .	$\bar{\psi}$	$n$ .	$k$ .	$\bar{A}$ .	$\bar{\psi}$ .	$n$ .	$k$ .	$\bar{A}$ .	$\bar{\psi}$	$n$ .	$k$ .
$^{\circ}$	$^{\circ}$	$^{\circ}$			$^{\circ}$	$^{\circ}$			$^{\circ}$	$^{\circ}$		
95	49.16	42.81	1.67	3.04	45.80	44.02	1.36	3.47	39.5	45.3	1.06	3.91
90	88.86	39.78	1.69	2.96	85.76	40.87	1.39	3.31	73.7	42.9	1.00	4.04
80	136.88	41.01	1.68	2.90	133.60	41.66	1.47	3.22	120.5	42.2	1.02	3.79
70	160.51	43.68	1.68	2.89	158.05	44.06	1.45	3.25	147.4	44.3	0.93	4.06
60	174.15	45.65	1.69	2.87	172.90	45.94	1.43	3.22	166.0	46.0	0.97	4.03
50	183.60	47.06	1.76	2.75	182.47	47.23	(1.84)	(3.12)	177.9	47.0	1.03	3.60
40	190.06	48.05	(1.87)	(2.53)	189.66	48.13	(1.79)	(2.75)	187.5	48.1	(0.97)	(3.27)
Mean values of $n$ and $k$				2.90	—	—	1.42	3.30	—	—	1.00	3.91

in agreement with those obtained for the solid metals.\* The mean values of the optical constants obtained from them are shown in Table IV.

Table IV.

$\lambda$ .	$n$ .	$k$ .	$\Phi$ .	$\Psi$ .
A.			°	°
5780	1.77	2.8	88.64	39.23
5400	1.61	3.1	88.17	40.04
4358	1.03	4.0	85.48	42.11

The  $\Delta$ - and  $\psi$ -values in the tables are deduced from readings of the polarimeter, each of which is the mean of from five to ten settings. The polarizer readings are probably accurate to less than  $0.1^\circ$ , and those of the analyser to  $0.05^\circ$ ; hence the error in  $\Delta$  and  $\psi$  are  $\pm 0.2^\circ$  and  $\pm 0.05^\circ$ , and so those in  $n$  are  $\pm 0.02$ . The values for the various series set out in Tables I, II, and III differ from each other by amounts larger than these errors, indicating that the surfaces are not all identical (see below). In any one series, however, the values quoted probably express the state of the surface observed to within these limits of error.

The  $n$  values of any one run do not agree to within  $\pm 0.02$ ; their steady increase is discussed below, and in the tables the mean of those values not enclosed in brackets is taken as the value of  $n$ . The rejected values seem far enough removed from the mean values to justify their rejection. Judging from the figures of which they are the mean, these values of  $n$  are probably accurate to  $\pm 0.05$ , and the corresponding values of  $k$  to  $\pm 0.1$ .

#### Discussion.

In Table V are given the values of  $n$  and  $k$  which have been obtained up to date. Owing to the large variety of wave-lengths employed the  $n$  values are plotted in fig. 2. Whereas observations can be made to an accuracy certainly within  $\pm 0.2^\circ$  in  $\Delta$  (representing  $\pm 0.02$  in  $n$ ), the published values of  $n$  vary by as much as 0.4. Certain differences are to be expected, as the various workers used mercury surfaces in contact with different media; Meier† used a vertical surface in a cell on the face of a "uvio" prism, Drude‡ and O'Brien§

\* Cf. Schulz, 'Wien's Handb. Experimental physik,' vol. 18, p. 554, Leipzig (1928).

† 'Ann. Physik,' vol. 31, p. 1031 (1910).

‡ 'Ann. Physik,' vol. 39, p. 531 (1890).

§ 'Phys. Rev.,' ser. 2, vol. 27, p. 93 (1926).



worked in air, Haak\*, Ellerbroek† and Reeser‡ worked *in vacuo*, and the authors worked in carefully cleaned nitrogen. Nevertheless, the differences observed are very large, 0.4 units in  $n$  representing at least 4° in  $\Delta$  at 90° incidence, and, as can be shown from equation (5.1), this would correspond to a film (for a fatty acid) 15 Å. thick. Hence these differences must be regarded, at least in part, as due to contamination.

Table V—Summary of published values of  $n$  and  $k$ .

Observer	$\lambda$ .	$n$	$k$ .	Observer.	$\lambda$ .	$n$ .	$k$ .
	Å.				Å.		
Meier	3982	0.921	3.438	Ellerbroek	5580	1.38	3.58
O'Brien	4017	0.79	4.31	Reeser	5580	1.592	3.00
O'Brien	1358	0.88	3.94	Reeser	5780	1.639	3.01
Reeser	4358	0.905	3.772	The Authors	5780	1.77	2.84
The Authors	4358	1.03	3.95	Haak	5893	1.68	2.98
Meier	4413	1.011	3.387	Meier	5893	1.624	2.713
Meier	4678	1.149	3.207	Drude	5893	1.73	2.87
Reeser	4820	1.198	3.552	Meier	6300	1.749	2.733
Meier	5083	1.312	2.985	Drude	6300	1.87	2.78
Reeser	5460	1.538	3.05	Reeser	6370	2.076	2.61
The Authors	5460	1.61	3.10	Reeser	6690	2.218	2.49

In fig. 2 it will be noticed that the values of Reeser lie consistently below those of the authors, whereas those of Drude and O'Brien lie yet lower. These latter can definitely be considered to be values for contaminated surfaces, since the mercury was exposed to air. Burdon§ has shown that very clean mercury in clean dry air keeps its surface uncontaminated, and the surface-tension remains constant. In ordinary air, however, the surface-tension falls, indicating a contamination of the surface, whereas if the mercury be impure a visible scum forms.||,¶ Thus, it is to be expected that the  $n$ -values of Drude and O'Brien will be lower than the true values of  $n$ .

Meier's values are of interest, as his surface was protected from contamination. His  $n$ -values lie on a curve which cuts the authors' curve sharply. According to a previous paper of one of the authors, however,\*\* this can be interpreted as due to an adsorbed film of gas on the glass, and a film of impurity on the

\* 'Arch. Néerland,' vol. 6, p. 198 (1923).

† 'Arch. Néerland,' vol. 10, p. 42 (1927).

‡ 'Arch. Néerland,' vol. 6, p. 225 (1923).

§ 'Trans. Faraday Soc.,' vol. 28, p. 886 (1932).

|| Burdon, 'Proc. Phys. Soc., Lond.,' vol. 38, p. 148 (1926).

¶ Burdon and Oliphant, 'Trans. Faraday Soc.,' vol. 23, p. 205 (1927).

\*\* Tronstad, 'K. norske Vidensk. Selsk. Skr.,' No. 1, pp. 38, 207 (1931).

mercury, with refractive indices respectively lower and higher than that of the glass.

Excepting those of Meier, the values of Reeser lie closest to those of the authors; that is to say, to those in Table I. In Tables II and III the authors' values fall off until they are near to, or slightly below, those of Reeser. Reeser cleaned his surfaces by pouring some excess mercury away, but it was observed by Burdon\* that impurities cannot be completely removed in this manner. This was also observed by the authors with the grey contamination referred to above. Considering this fact, and also that Tables II and III are known to apply to slightly contaminated surfaces, it seems probable that Reeser's

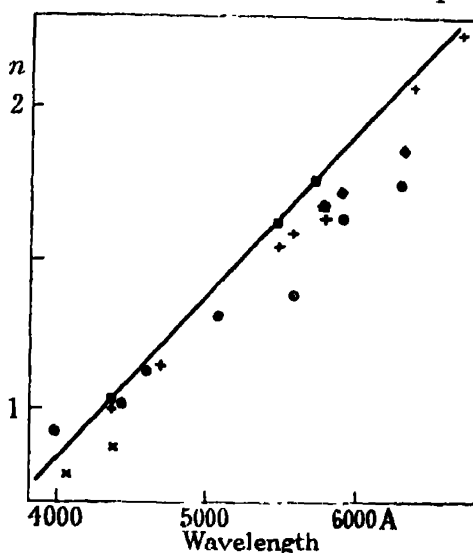


FIG. 2.—◇ Drude; ⊙ Ellerbrook; \* Haak; ● Meier; × O'Brien; + Reeser; □ Authors.

surfaces were also slightly contaminated. In this event the values in Table I may be considered the nearest approach to the true values yet obtained. It is necessary to emphasize the great importance of the accurate determination of the angle of incidence. It can be shown from equation (1.1) that the errors caused by small errors in the angle of incidence are very large, and discrepancies may be due to neglect of exact determination of this angle.

It will be observed that there is a steady increase in  $n$  and a decrease in  $k$  with decreasing  $\phi$ . There are no published data with which to compare this for mercury, and sometimes  $\phi$  is not specified. However, it is probable that it is a real effect. First, the changes in  $n$  and  $k$  as  $\phi$  varies are greater than

\* 'Proc. Phys. Soc. Lond.,' vol. 38, p. 148 (1926).

the experimental error. Secondly, the increase in  $n$  is less in Tables II and III than in Table I. Thirdly, values of  $n$  have been calculated from  $\Delta$ - and  $\psi$ -values of surfaces on which films of myristic acid had been placed (see Part II). These values of  $n$ , though less regular than those in Table I, are grouped round a mean value, and show much less tendency towards any steady increase. This suggests that for clean surfaces the values of  $n$  calculated from Drude's approximate equations are not independent of  $\phi$ , and consequently the divergencies between the  $n$  and  $k$  values of the different authors may also in fact be due to observations at different angles of incidence.

A further source of error is introduced because this method of determining the optical constants of a medium assumes a perfectly sharp change from the medium to the gaseous phase; this assumption cannot be applied to mercury. Even if the mercury were quite clean, and the molecular thermal agitation at the surface very small, it is still possible that mercury contains more than one species of molecule (*e.g.*, that it is a mixture of Hg and Hg<sub>2</sub>) and that there is a Gibbs layer at the surface with a composition different from that of the bulk. This would cause a sharp change in properties through a depth, possibly smaller than the wave-length of light, but probably greater than one Angström unit. In an absorbing medium the reflected ray penetrates about half a wave-length, and any such change in properties would affect the light in the same way as an absorbing layer. Such a layer, more than one or two Angström units deep, would be detectable by the optical method used in this work, so that liquid mercury could never present an ideal surface as, for instance, a cleavage plane of a solid crystal would. This being so, the accuracy to which  $n$  and  $k$  for bulk mercury can be found by this method is limited, and may be indicated by the divergence of the values of  $n$  for varying values of  $\phi$ .

The authors wish to express their thanks to Professor E. K. Rideal, F.R.S., for his interest in the work, to Mr. L. Saggers for many valuable suggestions with regard to the instrumental part and to the Fondent "Nationalgaven til Chr. Michelsen," Bergen, Norway, for a fellowship to one of us (L.T.) and to the Department of Scientific and Industrial Research for a maintenance grant (C. G. P. F.).

#### *Summary.*

The values of the optical constants of mercury have been measured for  $\lambda\lambda$  4350, 5460, and 5780 Å.

The values are discussed in relation to previously published values for wave-lengths in the visible region of the spectrum.

---

*An Optical Examination of Thin Films. II.—The Behaviour of Thin Films of Fatty Acids on Mercury.*

By C. G. P. FEACHEM and LEIF TRONSTAD.

(The Laboratory of Colloid Science, Cambridge.)

(Communicated by E. K. Rideal, F.R.S.—Received December 22, 1933.)

*Introduction.*

The optical method of examining plane surfaces, described in I, p. 115, was applied to fatty acid films on mercury, in the hope that it would be possible to determine the optical properties of monolayers. The difficulties of making and keeping large clean surfaces of mercury are well known, but the optical method possesses this advantage over the surface-pressure and surface-potential methods, that it requires only a relatively small surface. Mercury possesses advantages over water as a subject for the optical method, in that it reflects most of the light that falls on it, and no light is reflected or scattered from the bottom of the liquid, as with transparent media.

An analogous optical method has been most successfully applied to the study of surfaces by Bouhet.\* He studied the Gibbs-layer on saturated solutions of the lower fatty acids, and found that it merged continuously, as he went up the series of acids, into the solid surface-films of the higher acids. Observations on films on mercury have been carried out by Schulman† and Fahir,‡ both working in currents of nitrogen. Schulman distilled mercury into a large dish, and found that films could readily be spread and that they caused large changes in the air-mercury interface potential. Fahir used a trough edged with sheets of cellophane, and cellophane barriers, and was able to measure surface-pressures, which he found to be about twice as great as the corresponding pressures on water. Neither of these workers, however, makes any claim to have had dry surfaces.

*Experimental Method.*

Observations were carried out in the chamber described in I, either in a vitreosil dish or a steel trough. The mercury was distilled *in vacuo* into the

\* 'Ann. Physik,' vol. 15, p. 5 (1932).

† *Unpublished work.*

‡ 'J. Chim. phys.,' vol. 27, p. 587 (1930).

bottom of a glass trap. By admitting nitrogen, it could be pushed up the centre tube of the trap, and delivered into the chamber by the device shown in fig. 1. The mercury flowed along the glass tube A, ending in the jet B. On the roof of the chamber was fixed a brass tube D, fitted by a ground joint to a further tube E, similarly fitted to a third tube F. The two ground joints were carefully made co-axial. The glass tube A, which had to be fixed in space, was sealed into the collar F with wax, but its jet B projected well into the tube E. Into this tube E was sealed the glass tube G surrounding the jet B, and projecting down into the chamber, where it had two right-angle bends, and

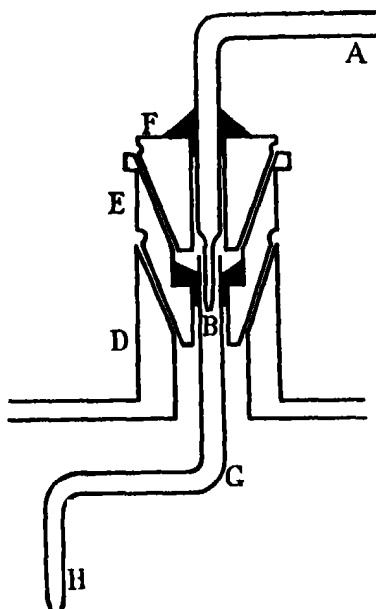


FIG. 1.

ended in the jet H. This jet could be swung by twisting the tube E, so that the mercury either fell into the dish, or clear of it, into a suitable receptacle. With this type of joint, the mercury touches nothing but glass. In the chamber was a vessel of phosphoric oxide to remove the last traces of water, and maintain the surface dry.

In some experiments, the mercury was allowed to fall into a vitreous dish. Attempts were made to clean the surfaces thus produced by allowing the mercury to overflow, but when dust particles could be seen on the surface, it was observed that they did not all go over the edge, a feature noted also by Burdon and Oliphant.\* A rectangular steel dish was also constructed with ground under-edges and steel barriers,

for the purpose of altering the size of

the surface. The barriers were held down on to the edges of the trough by spring-clips, and one of them could be moved by means of a screw, worked from outside the chamber. Sweeping the surface with these barriers made it appreciably cleaner, but owing to the large contact angle of mercury, which does not wet steel, it was not always possible to prevent leaks in the surface at the ends of the barriers. When the edges and the barriers were coated with cellophane the tendency to leak was diminished but not always eliminated.

Films having been prepared, the values of  $\Delta$  and  $\psi$  for the various surfaces were measured, as described in I, in the mercury green light ( $\lambda = 5460 \text{ \AA}$ ).

\* 'Trans. Faraday Soc.,' vol. 23, p. 205 (1927).

From these can be deduced the "thickness" and refractive index of the film, assuming it to be uniform, from equations (1.1), (2.1) of Part I, or the scattering indices of the molecules, from the equations of Strachan.\* It is clear that the optical "thickness" of a unilm may not be identical with the actual value and that "refractive index" is no more than an optical property, but comparisons of thicknesses of films might be expected to contrast various states of the film. This can be effected equally well by comparing the  $\Delta$ -values as in the following paragraphs. The "thickness" is proportional to  $(\Delta - \bar{\Delta})$ ; this latter quantity is called "the  $\Delta$ -value" of each film. 1°  $\Delta$ -value corresponds roughly to 3 Å.

### *Spreading of Films.*

When crystals or drops of fatty acids are placed on the surface of clean dry mercury they spread very slowly. With myristic and lauric acids it took one or two days before an appreciable belt of film had formed (about 1 mm. wide) wide enough to measure. The liquid acids were much less sluggish, caproic acid spreading out several millimetres in a few minutes. Oleic acid, on the other hand, spread at once and violently, as it does on water, and then lenses formed in the film. As well as spreading from bulk, the acids were also spread from solvents, including petrol-ether and chloroform. The redistilled petrol ether (A.R.) when tested by measuring the surface potentials on water left no residue which affected the potential difference. On mercury, however, it left a residue which had a  $\Delta$ -value of 0.3°, which was not altered by evacuating the chamber or by repeated redistillation, but which disappeared when the surface was left standing for several hours in the presence of phosphoric oxide. Thus it is probably due to traces of water which would not be removed by a Hyvac pump. From the experiments of Cassel and Salditt† we may conclude that water is strongly adsorbed on to mercury at very low vapour pressures. With chloroform, the contamination could be kept down as low as 0.2°. For this reason chloroform was more often used; this would not always spread to form a thick layer before it evaporated. Usually it formed a lens, but a thin layer appeared to spread from it, for when the lens had evaporated, the surface was covered with a uniform film. On dirty surfaces the lens either evaporated, leaving crystals behind, or rolled over the curved edge of the mercury.

\* 'Proc. Camb. Phil. Soc.,' vol. 29, p. 116 (1933).

† 'Z. phys. Chem.,' A, vol. 155, p. 321 (1931).

Myristic and lauric acids were spread from solvents. When the spreading was successful, it was observed that uniform surfaces were obtained when the extension of the film was greater than  $50 \text{ \AA}^2$  per molecule. These films had  $\Delta$ -values up to  $2^\circ$ . This, unfortunately, is only just outside the combined experimental errors, but it is possible to say that the more extended films had, on the whole, smaller  $\Delta$ -values, reaching to about  $0.9^\circ$  with a molecular area of  $650 \text{ \AA}^2$ .

When sufficient film material was applied to give films of specific areas less than about  $50 \text{ \AA}^2$  per molecule, patches of one or two additional phases were formed, thicker than the thin phase just described. Of these phases, only the thickest of all appeared to be ultimately stable. The lower acids, down to caproic, were also investigated, being spread both directly and from solvents, and the same three phases were observed.

(1) The solid phase. This consists of a solid film, having a  $\Delta$ -value of from  $10^\circ$  to  $15^\circ$ , depending on the chain length of the acid.

(2) A liquid phase, with a  $\Delta$ -value of  $5^\circ$  to  $6^\circ$  when in equilibrium with the solid phase. This phase can be expanded down to  $2^\circ$  in equilibrium with

(3) the expanded phase. For lauric acid, this phase terminates at an area at greatest compression of about  $50 \text{ \AA}^2$ , which is about the limit of the vapour expanded phase on water. This phase can be extended apparently indefinitely with ever-decreasing values for  $\Delta$ .

When composite films were made, they usually consisted of the expanded phase, and one or both of the condensed phases. In time, after some hours, all the material tended to pass into the solid phase. Patches of the liquid film vanished, and the size of individual patches of solid film increased. At the same time, the expanded film became thinner, so that finally after about a day no measurable  $\Delta$ -value could be obtained. This suggests that the final state of equilibrium would involve only the solid and the undetectable vapour phases, and that the liquid phase is metastable.

The  $\Delta$ -values, which are proportional to the thicknesses, of these three phases, are set out in Table I.

It will be noted that oleic acid (with a double bond between C 9 and C 10) has a  $\Delta$ -value which would fall into the series at about C 11, suggesting that in the films of oleic acid observed, the double bond was on the surface, having a chain of nine atoms in the air, and the rest of the molecule coiled near the surface, giving a nett thickness rather greater than that appropriate to C 9.

These films may be compared with those obtained by Bosworth\* by adsorption from the vapours of the lower fatty acids. He obtained a condensed phase with a molecular area of  $22 \text{ \AA}^2$  at its greatest extension, and an expanded phase whose molecular area increased with chain-length (for acid  $C_n$ ,  $A = [25 + 5(n - 1)] \text{ \AA}^2$ ). The films here described have a thick phase with a thickness which increases with chain-length, and so probably correspond with Bosworth's condensed phase, and a second phase, of constant thickness, presumably corresponding with Bosworth's second phase of regularly increasing area.

Table I.

Acid.	No. of C. atoms.	$\Delta$ for solid phase.	$\Delta$ for liquid phase at extreme compression.	$\Delta$ for expanded phase.
Caproic	6	—	$5.2^\circ$	—
Caprylic	8	10.4	—	0.8
Pelargonic	9	—	5.1	1.8
Lauric	12	11.9	5.6	0.3
Myristic	14	14.7	6.4	0.9
Oleic	18	11.5	—	0.0
	+ a double bond			

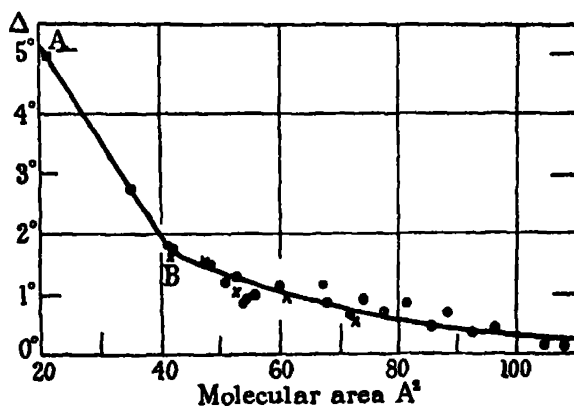


FIG. 2.

### Compression of Films.

In spite of the fact that the steel trough with barriers did not always give a surface free from leaks, it was possible to carry out a number of runs involving compression and re-expansion of films. The results for lauric acid are combined as a curve in fig. 2. Changes in the expanded film were always reversible,

\* 'Trans. Faraday Soc.,' vol. 28, p. 903 (1932).



and, unless the system leaked, this film could be compressed up to  $42 \text{ \AA}^2$ , (B), without changing its character. Beyond this point some films could not be taken at all; others went a little way, and the slope of the curve increased sharply, but all, sooner or later, separated into a solid and a thin expanded phase. One film was carried to the point A, at  $22 \text{ \AA}^2$  per molecule, and had reached a  $\Delta$ -value of  $5^\circ$ ; that is to say, the most compressed state of the liquid film. On further compression, patches of solid phase appeared, but after a small drop in the  $\Delta$ -value of the liquid phase, it rose again to  $5^\circ$ . The surface was thus covered with solid and liquid phase in equilibrium.

### *Nature of Films.*

It is, of course, possible to suggest that the film did not consist of lauric acid, but of mercury laurate, the formation of which from mercury and thick layers of acid has been observed by Trillat.\* However, only occasionally were changes in the film observed which could be so interpreted. Most of the films behaved consistently while under observation, so that if salt-formation occurred it must have occurred immediately. In addition, except for such water as was introduced with the film, the surfaces were quite dry, so that it is unlikely that salt-formation took place.

### *Deductions from the $\Delta$ - and $\psi$ -values.*

From the equations of Drude already referred to the thickness  $L$  and refractive index  $n_1$  of a uniform thin film can be deduced from the  $\Delta$ - and  $\psi$ -values, and the calculated thicknesses are set out in Table II.

Table II.

State of film.		Expanded.	Liquid.	Solid.
Area per molecule.		$45 \text{ \AA}^2$	$22 \text{ \AA}^2$	?
Acid.	Chain-length.			
	A.	A.	A.	A.
Caproic .. .. .	8	—	15	—
Caprylic .. . . .	11	—	—	30
Pelargonic .. . . .	12	—	15	—
Lauric .. . . .	16	5.4	17	44
Myristic .. . . .	19	5.8	22	47
Oleic .. . . .	Two chains, 12 A. long separated by a double bond		33	

\* 'J. Phys. Rad.,' vol. 10, p. 32 (1929).

From the figures for myristic and lauric acids, it might be suggested that the liquid phase consists of close packed orientated molecules, and the solid phase of a bimolecular leaflet. This, however, does not agree with the values for the shorter acids whose solid phases would have to be supposed to consist of three layers. In addition, it seems probable that the equations of Drude for a homogeneous, isotropic layer, would not give values of the "thickness" of a highly inhomogeneous and anisotropic unimolecular layer with any great accuracy, and that the calculated values are systematically too great. The liquid and solid phases may thus be regarded as corresponding to the two phases observed by Bosworth, as described above.

### Scattering Indices.

The equation found by Strachan\* connecting the scattering indices of the surface per unit area with the  $\Delta$ - and  $\psi$ -values, considers three such indices,  $\sigma_1$  and  $\sigma_2$  in the plane of the film, and  $\sigma_3$  normal to it. Assuming  $\sigma_1 = \sigma_2$ , and putting in the appropriate constants, this equation reduces to

$$\sigma_1 = -[3 \cdot 14 (\Delta - \bar{\Delta}) + 15 \cdot 4 (\psi - \bar{\psi})] 10^{-8} \text{ cm.} \quad (1.2)$$

$$\sigma_3 = -[0 \cdot 0764 (\Delta - \bar{\Delta}) + 0 \cdot 616 (\psi - \bar{\psi})] 10^{-8} \text{ cm.} \quad (2.2)$$

Knowing the extension of a film, the scattering index per molecule can be found. This quantity has the dimensions of a volume, and can be compared with the index for the same substance in bulk. Darwin† has shown that, to a first approximation, if  $n$  is the refractive index of a medium in any direction, the scattering index in that direction is given by

$$\sigma \left( \frac{1}{3} + \frac{1}{n^2 - 1} \right) = 1. \quad (3.2)$$

That is to say

$$\sigma = \frac{3}{n^2 + 2} \quad (4.2)$$

The refractive indices of the liquid fatty acids have been measured,‡ from which the  $\sigma$ -values can be found. Except for formic acid, they obey very closely the law

$$\Sigma = (0 \cdot 20 + 0 \cdot 225n) 10^{-22} \text{ c.c.,} \quad (5.2)$$

where  $\Sigma$  is the molecular scattering power and  $n$  the number of carbon atoms in the molecule.

\* 'Proc. Camb. Phil. Soc.,' vol. 29, p. 116 (1933).

† 'Trans. Camb. Phil. Soc.,' vol. 23, p. 337 (1924).

‡ 'Int. Critical Tables,' vol. 7, p. 37.

Data for the solid acids are not available except for lauric acid\* which is reported as having  $n = 1.48$ , double refraction small. Assuming that its density is 0.9, and that it is isotropic, this gives

$$\Sigma = 3.15 \cdot 10^{-22} \text{ c.c.}$$

whereas the law of equation (7) gives

$$\Sigma = 2.90 \cdot 10^{-22} \text{ c.c.}$$

Assuming a mean value of  $3.0 \cdot 10^{-22}$  c.c. for lauric acid, myristic acid has, by extrapolation,  $\Sigma = 3.5 \cdot 10^{-22}$  c.c.

The values calculated from observations are given in Table III.

Table III.—Scattering Indices for Molecule.

Acid.	Acid in bulk.	State of film.	Acid as film.	
			In plane of film.	Normal to film.
Myristic	$10^{-22}$ c.c. 3.5	Expanded Liquid	$10^{-22}$ c.c. 2.3	$10^{-24}$ c.c. 3.6
			2.8	4.1
Lauric	3.0	Liquid	1.9	2.3

Scattering indices cannot be evaluated for solid films, since the observed values of  $(\psi - \bar{\psi})$  are too large.

These results show that the scattering indices, in the plane of the film, of the molecules of the liquid and expanded phases are of the same order of magnitude as those of the corresponding molecules in bulk. The indices for vibrations normal to the film are about 100 times smaller. The packing of molecules in solid films is such that they cannot be regarded as independent scattering centres, and so do not give positive indices.

The authors wish to express their thanks to Professor E. K. Rideal, F.R.S., for his advice and assistance during this work, and to Dr. O. H. Wansbrough-Jones and the members and staff of this Department for their interest and co-operation.

#### *Summary.*

The behaviour of several fatty acids as films on mercury has been studied by the method of examining the surface in polarized light. It is found that :—

(1) The acids spread with extreme difficulty to give films, the shorter-chain acids spreading more readily than the longer.

\* 'Z. Krystallog.', vol. 67, p. 570 (1928).

(2) The films can exist in three phases, with molecular areas as follows :—

Solid, up to  $22 \text{ \AA}^2$ ; liquid, 22 to  $45 \text{ \AA}^2$ ; expanded, from  $45 \text{ \AA}^2$ .

Of these phases only the solid and expanded are ultimately stable.

(3) Homogeneous films of expanded phase can be prepared and compressed through the liquid phase, sometimes up to  $22 \text{ \AA}^2$  per molecule.

It is concluded that the method is amply sensitive to measure unimolecular and other thin films, but better suited to examine adsorption.

---

### *The Atomic Work Function of Tungsten for Potassium.*

By R. C. EVANS, B.A., B.Sc., Denman Baynes Research Student, Clare College, Cambridge.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received December 27, 1933.)

#### *Introduction.*

In a recent paper\* some experiments on the rate of evaporation of positive ions of the alkali metals from a hot tungsten surface were described, and, from the temperature coefficient of the evaporation rate the positive ion work function, the work required to remove a positive ion from the surface, was derived. In the experiments described here the method has been extended to investigate the rate of evaporation of potassium *atoms* from a tungsten surface, and the atomic work function has been measured.

#### *Theory.*

Consider a hot tungsten surface on which an atomic beam of potassium is incident, and let a negative field be maintained at the surface to prevent positive ions from evaporating. Under these conditions potassium can leave the surface only in the atomic state and the surface concentration will thus go on increasing until the rate of evaporation is equal to the supply in the incident beam, when an equilibrium state will be set up. The equilibrium surface concentration  $N_{a_s}$  will be given by†

$$N_{a_s} = Q/a,$$

\* Evans, 'Proc. Roy. Soc.,' A, vol. 139, p. 604 (1933).

† Evans, 'Proc. Camb. Phil. Soc.,' vol. 29, p. 161

where  $a$  is the atomic evaporation rate for unit surface concentration, and  $Q$  is the intensity of the incident beam. The corresponding ionic surface concentration is

$$N_{p_0} = fN_{a_0}/g = fQ/ag,$$

where  $f$  and  $g$  are quantities such that  $fN_{a_0}$  is the number of adsorbed atoms which are converted into ions and  $gN_{p_0}$  is the number of adsorbed ions which are converted into atoms on unit area of the surface per second. We shall, therefore, have

$$f/g = \frac{1}{2} \exp(-I_s e/kT),$$

where  $I_s$  is the surface ionization potential, and the numerical factor  $\frac{1}{2}$  arises from the different weights to be attributed to the potassium atom and ion.

If now the field at the surface be reversed, the excess concentration of adsorbed atoms and ions will evaporate until the new equilibrium is reached for which the evaporation in the atomic and ionic states together equals the arrival rate. The new equilibrium surface concentrations will thus satisfy the equation

$$aN_{a_0} + bN_{p_0} = Q,$$

where  $b$  is the ionic evaporation rate for unit surface concentration. In fact, these concentrations are\*

$$N_{a_0} = (b + g) Q/w, \quad N_{p_0} = fQ/w, \quad (1)$$

where

$$w = ab + ag + bf.$$

The total quantity of electricity which leaves the surface when the field is reversed is thus

$$N_{p_0} - N_{p_1} = bf(a + f) Q/agw = S_0, \text{ say.}$$

The experimental procedure outlined is precisely that adopted in the previous experiments, the decay of current to a collector surrounding the target on reversing the field being measured with an oscillograph. The records obtained were of the form of the portion ABC of fig. 1; the current to the collector, zero before reversing the field, assumed a large value immediately after reversal, and decayed exponentially with time to a limiting value  $bN_{p_1}$ ; corresponding to the new equilibrium surface concentration. The shaded area under the curve measures the value of  $S_0$  appropriate to the incident beam and target temperature employed.

\* Evans, 'Proc. Camb. Phil. Soc.,' *loc. cit.*

Now let us assume that at C the field is reversed again, and after a further  $t$  seconds, once more made positive so that we obtain the curve EFG. If  $t$  is very long compared with the half period of the evaporation curves, the curve EFG will be identical with ABC, and in particular the area beneath it will be the same. If, however,  $t$  be gradually decreased, there will come a time when the interval DE is too short for equilibrium conditions to have been attained, and the area under EFG will then be less than that under ABC, being, in fact, the concentration that has accumulated in time  $t$ . In this way, by making a series of reversals with smaller and smaller values of  $t$ , we shall be able to

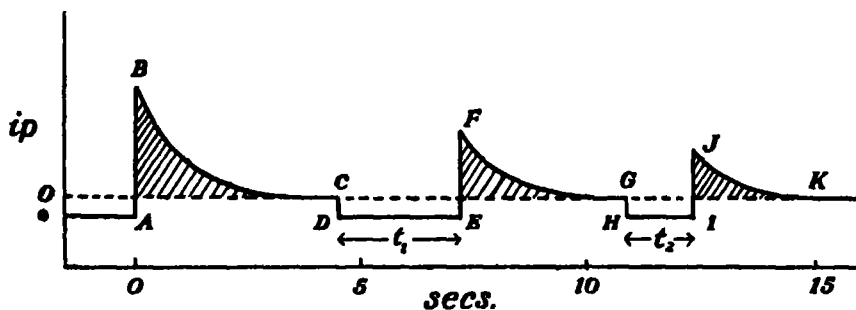


FIG. 1.

follow the growth of the surface concentration under the influence of the constant arrival rate, and the atomic evaporation rate by measuring the area under each of the exponential curves.

A simple application of the theory developed for the previous experiments shows that under the conditions which obtain here the ionic surface concentration will increase with time according to the equation

$$N_p = [1 - \{b(a + f)/w\} \exp \theta t] fQ/ag, \quad (2)$$

a quantity which reduces to  $N_p$  when  $t$  is zero, and to  $N_{p_\infty}$  when  $t$  is very large. Here  $\theta$  is given by

$$\theta = -ag/f = A \exp \{ -(\phi_a - I_a) e/kT \}, \quad (3)$$

where  $\phi_a$  is the atomic work function, and  $A$  is a quantity whose rate of variation with  $T$  is small compared with that of the exponential factor, so that over a small temperature range it may be considered constant. It follows from (1) and (2) that the total quantity of electricity leaving the surface when the field is reversed after it has been negative for a time  $t$  is given by

$$S = N_p - N_{p_\infty} = fQ \left[ \frac{1}{ag} - \frac{1}{w} \left\{ \frac{b(a + f)}{ag} \exp \theta t + 1 \right\} \right] = S_0 \{1 - \exp \theta t\}. \quad (4)$$

From (4) we see that from a set of records giving  $S$  as a function of  $T$  at any temperature we can find  $\theta$  at that temperature, and hence by making sets of observations at different temperatures we can find  $\theta$  as a function of  $T$  and so calculate  $\phi_s - I_s$  from (3).

#### *Apparatus.*

The apparatus was in all essential respects identical with that used in the positive ion experiments, and accordingly calls for no detailed description here. The general vacuum technique and the method adopted for introducing the

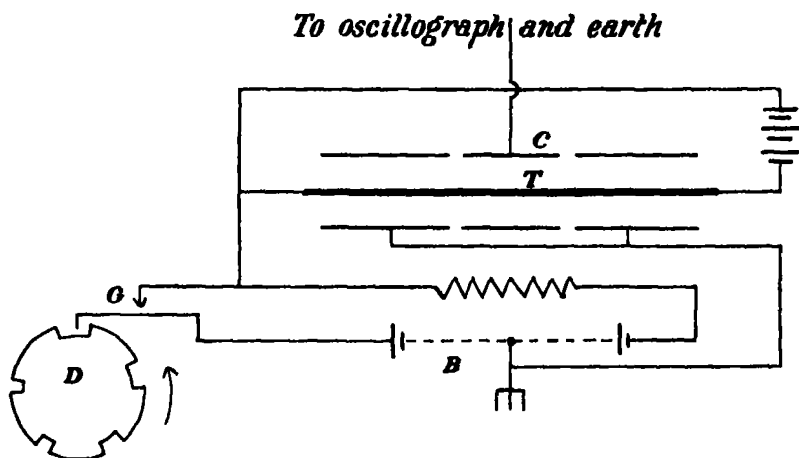


FIG. 2.

potassium were also the same. The essential electrical connections are shown diagrammatically in fig. 2. The apparatus was illustrated in detail in the previous paper.

It is sufficient here to state that a beam of potassium atoms from a gun was allowed to fall on the tungsten target  $T$  which took the form of a strip about  $60 \times 3 \times 0.05$  mm. This target was surrounded by a concentric cylindrical collector  $C$  constructed on the guard ring principle so that the measured currents came only from the central portion of the target, the temperature of which was sensibly uniform. The temperature of this part of the target was measured with a disappearing filament pyrometer sighted through a plane polished window.

The current from the target to the central portion of the collector was measured with the bridge system and oscillograph used in the previous experiments. The deflections of the oscillograph were recorded photographically on a moving strip of bromide paper.

The apparatus was constructed of pyrex glass and was outgassed by baking at  $480^{\circ}\text{C}$ . for an hour. The stills which contained the potassium projected outside the oven during this process. The lower part of the apparatus formed a mercury trap cooled in liquid air. Under the conditions which obtained during the experiments a McLeod gauge connected to the apparatus read zero. For further details reference may be made to the original paper.

The switching arrangement for reversing the field calls for special mention. The field between the target and collector was provided by the battery B, and was normally so directed that ions could not leave the target. By closing the contact G, however, this battery could be short circuited through a high resistance thus instantaneously reversing the field at the surface of the target. In order that equation (4) may be satisfied, it is necessary that the temperature of the target and the intensity of the incident beam remain constant throughout the time occupied by the several reversals of the field. This is best ensured by observing the exponential curves in a continuous sequence. The contact G was therefore operated automatically by a rotating disc D. In the edge of this disc five shallow slots were cut, the lengths of these slots measured along the circumference of the disc being in the ratios  $t_1 : t_2 : t_3$ , etc. The raised lugs between the slots were all of the same length, and several times the length of the longest slot. When the contact rested on a lug the field at the target was positive and positive ions could evaporate, while when it was in a slot the field was reversed. As the disc rotated the field was therefore alternately positive and negative, the periods for which it was positive being all the same and sufficiently long for the exponential curves to have decayed sensibly to zero, while consecutive periods  $t_1, t_2$ , etc., for which the field was negative were shorter and shorter. The rate of rotation of the disc could be varied between wide limits so that the times  $t_1$ , etc., could be altered while their ratio remained the same.

#### *Experimental Procedure.*

The experiments were performed in the following manner :—

The target was flashed to about  $2700^{\circ}\text{K}$ . for a few seconds and then allowed to cool to the experimental temperature  $T$  of about  $1000^{\circ}\text{K}$ . After 120 seconds, during which time the field at the target was kept negative, the camera and switch were started, and when the contact reached the first lug of the disc the field was reversed and the curve ABC recorded. After a time measured by AD, which was always long compared with the half period of the exponential curve, the contact fell into the first slot and the field was thus reversed for a time DE after which it once again became positive and the curve



EFG was traced. This procedure was continued until the switching mechanism had made one complete revolution, giving six curves in all.

A time base was recorded on the paper by imposing an A.C. wave of standard frequency on the zero line after making the experiment. The temperature of the target was measured within a few seconds of making the experiment, although it was found that the temperature remained very steady for long periods of time.

### *Results.*

The photographic records, each of the general form of fig. 1, but containing six, not three, curves, were analysed by plotting the logarithms of the deflections for each separate curve, measured from OK as zero, against the time. A straight line was always obtained indicating that to within the limits of measurement the curves BC, FG, etc., were exponential. The time constants of all the curves were not, however, quite the same, but increased slightly with decreasing length of exposure to the beam. This is in agreement with the observations made in the previous paper where it was shown that the time constant of the curves only tends to a limit after a considerable exposure to the beam. Since, however, we are here concerned only with the area under the curves this variation is immaterial. The time constants of the first curve of each set were in good agreement with those found in the positive ion work function experiments.

The area under each curve was calculated from the maximum ordinate and the time constant. In this way a series of areas  $S_0, S_1$ , etc., was found, the first being that corresponding to equilibrium conditions on the surface reached after long exposure, the others representing surface concentrations attained after  $t_1, t_2$ , etc., seconds. In some experiments the switching mechanism was arranged to give two records corresponding to long exposure, hence giving two values of  $S_0$ . The two values in all cases agreed to within about 5%, indicating that the intensity of the incident beam remained sensibly constant over the length of time required to make the six records constituting any one set. In other experiments the sequence of times  $t_1, t_2$ , etc., was altered so that they no longer formed a continuously decreasing sequence. The variation of  $S$  with  $t$  was found to be unaffected by this procedure.

Fig. 3 shows the result of analysing a typical set of curves for 1051° K. The curve WX illustrates how  $S$  changes with  $t$ , finally tending to the limit  $S_0$ , while the line YZ was obtained by plotting  $\log_{10}(S_0 - S)$  against  $t$ . The fact that the points so obtained lie on a straight line justifies an equation of the

form of (4). The slope of YZ is a measure of  $\theta$  at the target temperature used.

By making measurements at different temperatures a series of values of  $\theta$  was obtained, and a second graph was made by plotting  $\log_{10} \theta$  against  $1/T$ . This graph is reproduced in fig. 4, and it will be seen that no point lies farther from the straight line drawn than a distance corresponding to an error in temperature measurement of about  $4^\circ \text{C.}$ , which was found to be approximately the accuracy with which pyrometer readings could be reproduced.

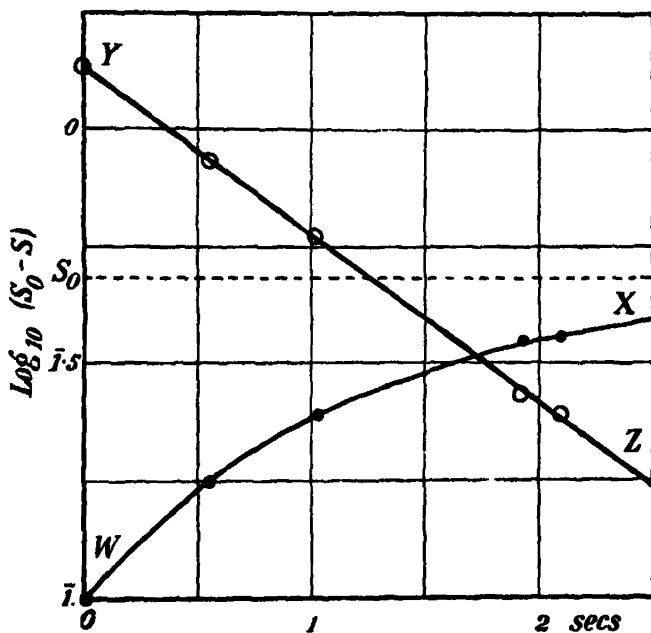


FIG. 3.

From (3) the slope of the line PQ is

$$-0.4343 \times (\phi_a - I_a) e/k = -5040 (\phi_a - I_a)$$

if  $(\phi_a - I_a)$  is measured in volts. The experimentally observed slope is  $-1.41 \times 10^4$ , whence  $(\phi_a - I_a) = 2.8_0$  volts.

### Discussion.

The general treatment adopted both here and in the discussion of the positive ion evaporation experiments has followed the lines of that due to Becker,\*

\* 'Phys. Rev.,' vol. 28, p. 341 (1926); 'Trans. Amer. Electrochem. Soc.,' vol. 55, p. 153 (1929).

who treats of atoms and ions co-existing on the surface, rather than that of Langmuir,\* who does not recognize a distinction between the adsorbed atoms and ions. The two treatments do, however, yield essentially the same results when we consider the atoms and ions evaporating from the surface. Thus in the present case where on Becker's treatment the number of adsorbed ions is very much larger than the number of adsorbed atoms, the evaporation of an atom will result in such a disturbance of the equilibrium on the surface that for

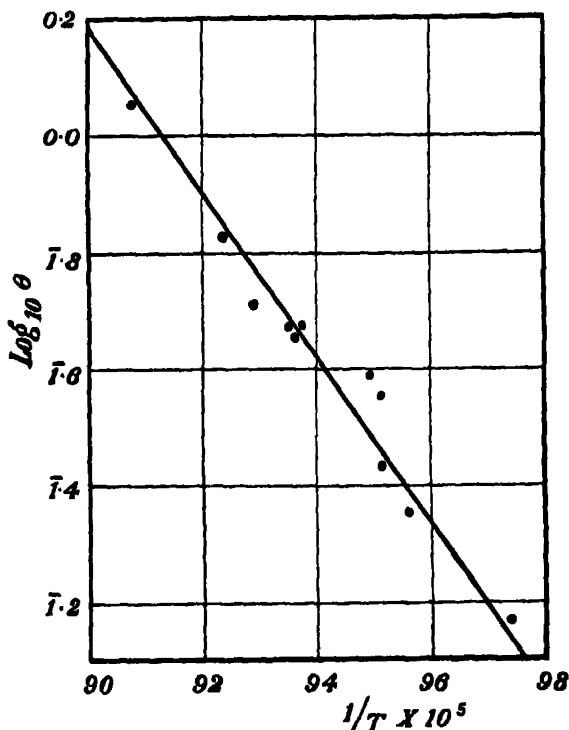


FIG. 4.

every atom which evaporates one adsorbed ion may be considered to change into the atomic state. If  $\phi_a$  is the actual work required to evaporate the atom, effectively the total work associated with the whole process of the atom evaporation is  $(\phi_a - I_a)$ , and it is this quantity that is measured by any experiment on the atomic evaporation rate. On Langmuir's treatment where the existence of a surface ionization potential is not recognized, this quantity would be considered as the atomic work function. If, however, we were able to calculate the work of evaporating a single atom in terms of the forces acting

\* 'J. Amer. Chem. Soc.,' vol. 54, p. 2798 (1932).

on an atom adsorbed on the tungsten surface we should expect to obtain  $\phi_a$  and not  $(\phi_a - I_e)$ . On the other hand, the evaporation of a single ion from the surface will not appreciably alter the surface equilibrium between atoms and ions, and therefore the quantity which we have designated  $\phi_p$  will be a true measure of the total work associated with the evaporation of an ion, and will be the quantity measured by an experiment on the ion evaporation rate. It is therefore to be expected that in this case agreement would be found between the value of the positive ion work function found experimentally and that calculated on some simple picture, such as that the ions are held to the surface by image force attraction. In the positive ion experiments such agreement was indeed found.\*

Although we are not in a position to make any such simple calculation on the values of  $\phi_a - I_e$ , we are able to apply a check to the value here obtained by considerations of the Born cycle. Let us suppose that we start with an atom on the surface. Then we can imagine the following reversible process. Let us ionize the atom on the surface, and let the atom and the electron be now evaporated independently and recombine in free space, the atom thus produced returning to the surface. Then we shall have,

$$\phi_p + \phi_a - I = \phi_a - I_e.$$

Inserting the value of 2.43 volts measured in the previous experiments for  $\phi_p$ , and the known values of  $\phi_a$  and  $I$ , we obtain for the left-hand side of the equation 2.73 volts.† This agrees to within the limits of experimental error with the value 2.80 volts here measured for  $\phi_a - I_e$ , for on account of the very narrow temperature range over which evaporation rates could be conveniently measured, the values of  $\phi_p$  and of  $\phi_a - I_e$  found experimentally cannot be considered accurate to more than about 5%.

Quite recently, Taylor and Langmuir‡ have published an account of some experiments on the evaporation of caesium from a hot tungsten surface in which they find a value of  $\phi_p$  of 1.91 volts when the adsorbed film is very dilute. This is in very satisfactory agreement with the value 1.81 volts found in the present author's positive ion experiments. The value of the atomic work function measured under the same conditions of surface concentration is given as 2.83 volts, while the value required to close the Born cycle is

\* Evans, 'Proc. Roy. Soc.,' *loc. cit.*

† The value 2.64 volts quoted in the previous paper was based on a value of  $\phi_p$  of 4.53 volts. The accepted value now appears to be 4.62 volts.

‡ 'Phys. Rev.,' vol. 44, p. 423 (1933).

2.65 volts. The difference is probably within the limits of experimental

It is a great pleasure to me to express my thanks to Professor Lord Rutherford, F.R.S., and to Dr. J. Chadwick, F.R.S., for their interest and advice, and also to Dr. M. L. E. Oliphant and to Dr. P. B. Moon for many valuable suggestions and discussions. I have also to express my thanks to the Department of Scientific and Industrial Research for a maintenance grant.

In the experiments here described, the method developed for the measurement of the rate of evaporation of *ions* of the alkali metals from a hot tungsten surface is extended to measure the atomic evaporation rate under the same conditions of very small ( $< 1\%$ ) surface concentration. By measuring at different temperatures the rate of change of surface concentration of the adsorbed ions under the influence of a constant arrival rate when *atoms* only are allowed to evaporate, a value of 2.80 electron volts was found for the work required to evaporate a potassium atom from the surface. This value taken in conjunction with the previously measured value of 2.43 volts for the positive ion work function is in very satisfactory agreement with the value 2.73 volts required to close the Born cycle.

---

## *On the Rotation of the Plane of Polarization of Long Radio Waves.*

By A. L. GREEN, M.Sc., and GEOFFREY BUILDER, Ph.D., F.Inst.P.

(Communicated by T. H. Laby, F.R.S.—Received December 27, 1933.)

### 1. *Introduction.*

The propagation of long\* waves at great distances has been studied in great detail by Austin,† by Round, Eckersley, Tremellen, and Lunnon,‡ by Espenschied, Anderson, and Bailey,§ and by Yokoyama and Nakai.|| Thus the propagational characteristics of long waves, considered merely as channels of communication, are well known. There have remained, however, a number of unexplained directional effects of which one might quote as an example the difficulty, noticed by Round, Eckersley, Tremellen, and Lunnon, of receiving signals whose great circle path traversed the earth's magnetic poles.

Pronounced directional effects at short distances have been reported by Naismith.¶ He found that the intensity of the space wave from GKB, Northolt, was approximately twice as great at Manchester as at Exeter although both receiving points are at the same distance from the transmitter. At Manchester the receiver was north of the sender, while at Exeter the direction was west.

Unfortunately the majority of these long-wave observations have been purely measurements of electric intensity in the vertical plane containing sender and receiver. The experimental results, by which one might hope to check any proposed theoretical explanation of the effects observed, are therefore ambiguous in the sense that too many variable quantities are involved. It is otherwise, however, with Hollingworth's\*\* exact measurements of polarization. By using an aerial system of two crossed loops and by measuring the c.m.f.'s in each of them as well as the phase angle between the two signals,

\* I.e., of the order 10,000 metres wave-length.

† Austin's publications are too numerous to mention in detail. Beginning with 'Bull. Bur. Stand.,' vol. 11, p. 69 (1914), there have been many papers in recent years in 'Proc. Inst. Rad. Eng.'

‡ 'J. Inst. Elec. Eng.,' vol. 63, p. 933 (1925).

§ 'Bell Syst. Tech. J.,' vol. 4, p. 459 (1925).

|| 'Proc. Inst. Rad. Eng.,' vol. 17, p. 1240 (1929).

¶ 'J. Inst. Elec. Eng.,' vol. 69, p. 875 (1931).

\*\* 'Proc. Roy. Soc.,' A, vol. 119, p. 444 (1928). Observations were made at Slough, distant 396 kilometres from the transmitter at Ste. Assise, wave-length 14,350 metres.

he has been able to separate the effects of ground and space waves. His discovery of a rotation of the plane of polarization in the downcoming wave occurring as a characteristic feature of the sunrise and sunset periods is summarized in figs. 1 and 2.

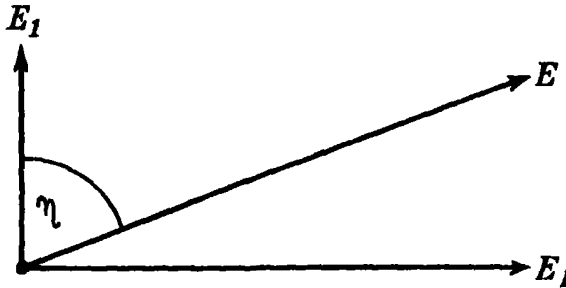


FIG. 1.—Electric forces in the downcoming wave, which is assumed to be travelling away from the observer and at right-angles to the plane of the diagram. For long waves, the normal and abnormal components  $E_1$  and  $E'_1$  are nearly in phase, so that their resultant  $E$  is a quasi-stationary vector whose plane of polarization is rotated through an angle  $\gamma$  from the normal position of  $E_1$ .

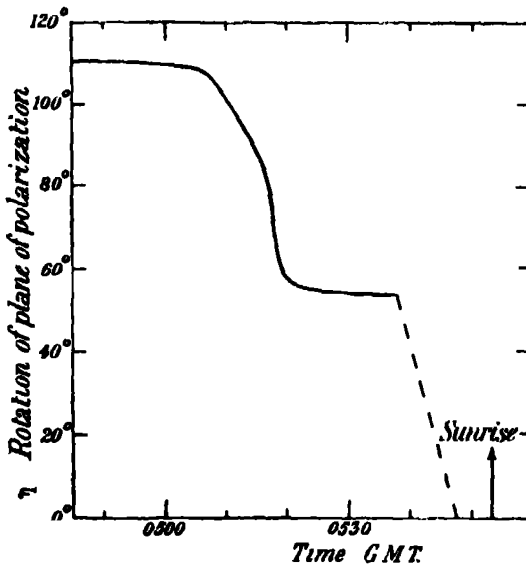


FIG. 2.—Experimental values of the rotation of the plane of polarization of the downcoming wave, as observed by Hollingworth for the morning of October 1, 1927. After 0540 G.M.T. the amplitude of the space wave appears to have been too small for satisfactory measurements of  $\gamma$ .

In fig. 1 the direction of propagation of the downcoming wave is assumed to be perpendicular to the plane of the diagram and away from the reader. The normal component of electric force, i.e., that contained in the vertical plane

which includes the sender and receiver, is  $E_1 \cdot \sin(pt + \phi)$ ,  $p$  being the angular frequency of transmission and  $\phi$  the phase angle with respect to the ground wave. The abnormal component, which is parallel to the ground, is

$$E'_1 \cdot \sin(pt + \phi + \xi),$$

where  $\xi$  is the phase difference between the electric components. Hollingworth's experimental results all agreed with his supposition that  $\xi$  was so small that the normal and abnormal components could be combined into a single equivalent force  $E$ , rotated through an angle  $\eta$  from the normal position of  $E_1$ . The orientation of this quasi-stationary vector was found to undergo definite cycles, a typical example\* of which is shown in fig. 2 for the sunrise period for October 1, 1927.

The main features of the observed rotation of the plane of polarization of the electric force  $E$  were :—

- (a) During the night the orientation of the plane of polarization was quite steady, the angle  $\eta$  being about  $110^\circ$ .
- (b) Beginning at 45 minutes before sunrise and lasting for about 10 minutes, there was an abrupt reduction in  $\eta$ , until a second steady state was reached at an angle of about  $50^\circ$  to  $60^\circ$ .
- (c) The second steady state lasted throughout the rest of the sunrise period during which measurements were possible. After 0540 G.M.T. on this morning the intensity of the space wave was small and the remainder of the curve, shown dotted, was filled in on the assumption that past experience had shown that the rotation was small during the day in summer. It would seem, however, that the dotted part of the curve should be disregarded since it was the intensity of the downcoming wave and not necessarily its rotation which was small during the day.

Characteristic sunset and sunrise cycles of field-intensity in the normally polarized component of electric force at the receiver, which are closely correlated with the rotation of the plane of polarization, have also been observed by Naismith and by Namba.† Nevertheless, the causes of these regularly occurring cycles of variation in the plane of polarization of the space wave have not yet, so far as we are aware, been elucidated. The present paper,

\* Copied from fig. 6 of Hollingworth's paper (*loc. cit.*).

† 'Proc. Inst. Rad. Eng.,' vol. 19, p. 1088 (1931). The receiving system used by Namba appears to be open to objections which would make it difficult to accept his observations of both right- and left-handed senses of rotation of the plane of polarization.



therefore, attempts to show that the technique of the magneto-ionic ray theory,\* which has been developed chiefly for short and medium wave propagation, is also adequate for the explanation of many of the apparent anomalies of long wave.

## 2. *Application of the Magneto-Ionic Theory to Long-Wave Propagation.*

(i) *Procedure.*—In what follows we shall use the methods advocated by Baker and Green.† These authors, it is true, were mainly concerned with the application of the magneto-ionic theory to medium-wave‡ propagation, but they give expressions for the limiting polarizations of the ordinary and extraordinary components§ of the downcoming wave which can readily be interpreted for use with long waves.

The procedure adopted here will be to calculate the limiting polarizations of both ordinary and extraordinary downcoming waves and then to combine them with suitable relative phases and amplitudes into a single space wave having, as in fig. 1, normal and abnormal components of electric force,  $E_1$  and  $E'_1$ . If, as the long-wave experimental results show, these components can be further combined into a single force  $E$  which is not rotating periodically, then the conditions which must hold will give a relation between (a) the experimental values of  $\eta$ , the angular rotation of the plane of polarization of  $E$  from its "normal" position,|| and (b) the relative phases and amplitudes of the two downcoming waves.

It will further be necessary to form estimates of the differential absorption experienced by the two space waves, owing to loss of energy in the ionosphere by electronic collisions, in order to see whether the amplitudes assumed for them are consistent with their probable attenuation factors.

\* Cf. Appleton, 'J. Inst. Elec. Eng.' vol. 71, p. 642 (1932).

† "Rad. Res. Board Rep. No. 3, Australian Commonwealth Council Sci. Ind. Res.," Bull. 60, Melbourne (1932); and 'Proc. Inst. Rad. Eng.,' vol. 21, p. 1103 (1933).

‡ I.e., transmission frequencies of about one-half of the gyroscopic frequency of electrons moving under the influence of the earth's magnetic field. At Slough the critical wavelength is 227 metres.

§ The ionosphere, as an aeolotropic medium, splits an incident wave into ordinary and extraordinary components, the former being that which is, for transmission frequencies not very far removed from the critical frequency, relatively little influenced by the presence of the earth's magnetic field.

|| The "normal" orientation of  $E$  should perhaps be defined as its steady day value, but this is different for each position of the receiving station. It is better, therefore, to follow Hollingworth and place the zero position of the plane of polarization of the downcoming wave in coincidence with that of the ground wave, i.e., the vertical plane containing sender and receiver.

(ii) *Relation between the Rotation of the Plane of Polarization and the Characteristics of the two Downcoming Waves.*—The ratios of axes of the ellipses of polarization in the ordinary and extraordinary downcoming waves, if we neglect the effect of collisions between electrons and other particles in the ionosphere, are given by

$$R_a^2 + \frac{l^2}{n} \cdot q \cdot R_a - 1 = 0, \quad (1)$$

where

$$l = \sin \theta, \quad n = \cos \theta, \quad q = p_H/p. \quad (2)$$

The roots  $R_a$  and  $R'_a$  are the ratios of axes of the two ellipses,  $\theta$  is the angle between the direction of propagation of the downcoming wave and the direction of the earth's magnetic field, and  $p_H$  is the electronic spiral angular frequency, having a value of about  $8 \times 10^6$  radians per second for English conditions.

In the ordinary downcoming wave the orientation of the ellipse is such that the major axis is in the plane containing the directions of the earth's field and of propagation. The principal components of electric force, namely, those along the major and minor axes of the ellipse, are in space and time quadrature. Hence, in the transfer of the ellipse to that which is measured at the ground, it is necessary to introduce the quantities  $R_o = E'_1/E_1$ ,  $\xi$  the angular phase difference between  $E'_1$  and  $E_1$ , and  $S$  the angle through which the ellipse has to be turned. There are corresponding quantities for the extraordinary wave,  $R'_o$  and  $\xi'$ , and the relations which we shall need are,

$$\cot \xi = \frac{(1 - R_a^2)}{2R_a} \cdot \sin 2S \quad (3)$$

$$R_o^2 = \frac{1 - \rho}{1 + \rho}, \quad (4)$$

where

$$\rho = \frac{(1 - R_a^2)}{(1 + R_a^2)} \cdot \cos 2S \quad (5)$$

$$\cos S = \frac{\cos i \cdot \cos \theta - \sin \delta}{\sin i \cdot \sin \theta} \quad (6)$$

$$\cos \theta = \cos i \cdot \sin \delta - \sin i \cdot \cos \delta \cdot \cos \gamma, \quad (7)$$

also

$$R_o \cdot R'_o = 1, \quad (8)$$

and

$$\xi' = \xi + \pi, \quad (9)$$

where  $\delta$  is the dip angle of the earth's magnetic field,  $\gamma$  the magnetic bearing of the sender from the receiver, and  $i$  the angle of incidence at the ground of the downcoming wave.

From these expressions it is possible to calculate  $R_o$ ,  $R'_o$ ,  $\xi$ , and  $\xi'$ . Let us now use  $E_1$  and  $E'_1$  as the normal and abnormal components of electric force, as measured at the ground, of the *combined* waves. We then have

$$\left. \begin{aligned} E_1 &= A_1 \cdot \sin(pt + \phi_1) + \frac{A_2}{R'_o} \cdot \sin(pt + \phi_2 - \xi') \\ E'_1 &= A_1 \cdot R_o \cdot \sin(pt + \phi_1 + \xi) + A_2 \cdot \sin(pt + \phi_2) \end{aligned} \right\}, \quad (10)$$

where  $\phi_1$  and  $\phi_2$  are the phases of the ordinary and extraordinary waves with respect to the ground wave, and where  $A = A_1/A_2$  may be defined as the ratio of electric intensities in the two downcoming waves.

Alternatively we may write

$$\left. \begin{aligned} E_1 &= \sqrt{A_1^2 + A_2^2 \cdot R_o^2 - 2A_1A_2R_o \cdot \cos(\phi_2 - \phi_1 - \xi)} \cdot \sin(pt + Z_n) \\ E'_1 &= \sqrt{A_1^2 \cdot R_o^2 + A_2^2 + 2A_1A_2R_o \cdot \cos(\phi_2 - \phi_1 - \xi)} \cdot \sin(pt + Z_a) \end{aligned} \right\}, \quad (11)$$

where

$$\left. \begin{aligned} \tan Z_n &= \frac{A_1 \cdot \sin \phi_1 - A_2 \cdot R_o \cdot \sin(\phi_2 - \xi)}{A_1 \cdot \cos \phi_1 - A_2 \cdot R_o \cdot \cos(\phi_2 - \xi)} \\ \tan Z_a &= \frac{A_1 \cdot R_o \cdot \sin(\phi_1 + \xi) + A_2 \cdot \sin \phi_2}{A_1 \cdot R_o \cdot \cos(\phi_1 + \xi) + A_2 \cdot \cos \phi_2} \end{aligned} \right\}. \quad (12)$$

On combining  $E_1$  and  $E'_1$  into the equivalent vector  $E$ , rotated through an angle  $\eta$  from  $E_1$ , we have

$$\left. \begin{aligned} E &= E_1 \cdot \cos \eta + E'_1 \cdot \sin \eta \\ 0 &= E_1 \cdot \sin \eta - E'_1 \cdot \cos \eta \end{aligned} \right\}, \quad (13)$$

whence

$$\tan \eta = E'_1/E_1, \quad (14)$$

and the condition that  $E$  should not rotate periodically is  $Z_n = Z_a$ . Now from (12) it follows that

$$\tan(Z_a - Z_n) = \frac{R_o \cdot \sin \xi \cdot \{A_1^2 - A_2^2 - 2A_1A_2R_o \cdot \cos(\phi_2 - \phi_1 - \xi)\} + A_1A_2(1 + R_o^2) \cdot \sin(\phi_2 - \phi_1)}{R_o \cdot \cos \xi \cdot \{A_1^2 - A_2^2 - 2A_1A_2R_o \cdot \cos(\phi_2 - \phi_1 - \xi)\} + A_1A_2(1 + R_o^2) \cdot \cos(\phi_2 - \phi_1)}. \quad (15)$$

Hence,

(a) if either  $A_1$  or  $A_2$  is zero, then  $Z_a - Z_n = \xi$ , which is the condition applicable\* in general to medium-wave propagation;

\* Cf. Appleton and Ratcliffe, 'Proc. Roy. Soc.,' A, vol. 117, p. 576 (1928); Green, 'Rad. Res. Board Rep. No. 2, Australian Commonwealth Council Sci. Ind. Res.," Bull. No. 59, p. 15, Melbourne (1932).

- (b) if  $R_e$  is zero, then  $Z_a - Z_n = \phi_2 - \phi_1$ , this being the condition applicable to propagation transverse to the earth's magnetic field, for all wave-lengths ;
- (c) if  $\xi = 90^\circ$ ,  $R_e = 1$ , and in addition  $A_1 = A_2$ , then  $Z_a = Z_n$ . Hollingworth\* finds that this condition has a general application to short-wave propagation ;
- (d) if  $\xi = 0$ , and in addition  $A_1 A_2 (1 + R_e^2) \sin(\phi_2 - \phi_1) = 0$ , then  $Z_a = Z_n$ .

Expression (15) has general application ; in the particular case of long-wave propagation we can safely assume that  $\xi$  is small for almost all conditions† that are likely to be met in practice, so that (d) above is relevant. It follows that the combined space wave is plane polarized when

- (a) either the ordinary or the extraordinary wave has zero intensity ; or
- (b) the two downcoming waves are in phase.

We may reject (a) as a special case of (b) and notice that the latter condition is consistent with Hollingworth's observation that night-to-day variations in the phase angle between the surface and space waves are small. It is to be inferred, therefore, that differential changes in phase between the components of the downcoming wave are of second order of smallness. Using (b) in (14), we have

$$\tan \eta = \frac{1 + A \cdot R_e}{A - R_e} \quad (16)$$

which is the required relation between the amount of rotation of the plane of polarization and the relative intensities of the ordinary and extraordinary rays.

(iii) *Differential Absorptions of the two Downcoming Waves.*—In the notation used by Baker and Green, the formula‡ for the complex refractive

\* 'J. Inst. Elec. Eng.,' vol. 72, p. 229 (1933).

† The exception is for transmission along the earth's field, when the polarizations of both downcoming rays are circular. For a wave-length of 10,000 metres, however, marked departures from linear polarization are only encountered when the angle between the direction of propagation and the earth's field is less than  $30^\circ$ . For these conditions the calculations of the next section show that  $A_1$  and  $A_2$  are nearly equal whence, from (c) above, the polarization of the combined wave is still plane.

‡ Hartree, 'Proc. Camb. Phil. Soc.,' vol. 27, p. 143 (1931), has suggested modifications to this formula which are designed to allow for local "polarization" effects in the medium. However, Baker and Green have demonstrated that such modifications cannot alter the limiting polarizations of either of the two downcoming rays. We have also been able to show that, for the conditions of long-wave propagation considered here, differential absorption of the two downcoming rays is similarly unaffected. Thus it is immaterial to the conclusions of the present paper whether the Hartree terms are included in the formula ; for brevity, they have been omitted.

index of the ionosphere is, after allowance has been made for electronic

$$(\mu - j\zeta)^2 = 1 - \frac{(1 - \epsilon_1) \{ \epsilon_2 - \frac{1}{2}l^2q'^2 \pm \sqrt{\frac{1}{4}l^4q'^4 + n^2 \cdot \epsilon_2^2 q'^2} \}}{\epsilon_1 + n^2 q'^2 (1 - \epsilon_1)}, \quad (17)$$

where

$$\begin{aligned} \epsilon_1 &= 1 + g' \\ \epsilon_2 &= 1 + g' - g'q'^2 \\ n' &= \frac{kq'}{p_H (q'^2 - 1)} \\ q' &= \frac{p_H}{p - j\nu} \\ k &= \frac{4\pi \cdot Ne^2}{mp} \end{aligned} \quad (18)$$

Also  $\mu$  is the refractive index,  $\zeta \cdot p/c$  the logarithmic attenuation factor per unit distance,  $N$  the number of electrons per cubic centimetre, of charge  $-e$ , mass  $m$  and frequency of collision  $\nu$ , and  $j$  has its usual operational significance.

Ordinarily the evaluation of  $\zeta$  is a difficult matter unless certain drastic assumptions are made. For long-wave propagation,  $p^2$  may be neglected in comparison with  $p_H^2$ . In addition we choose the condition\* that  $k \ll p$ . Expression (17) then reduces to

$$(\mu - j\zeta)^2 = 1 + g' \{1 - \frac{1}{2}l^2q'^2 (1 \pm 1)\} \quad (19)$$

if we make the further temporary assumption that  $n^2q'^2 \ll \frac{1}{4}l^4q'^4$ , which excludes propagation exactly along the earth's field. It can then be shown that,

$$\frac{\zeta_{ord}}{\zeta_{ext.}} = 1 + \frac{l^2 p_H^2}{p^2 + \nu^2} \quad (20)$$

and, for the special case of propagation along the earth's field,

$$\frac{\zeta_{ord.}}{\zeta_{ext.}} = 1 - \frac{4pp_H}{p_H^2 + \nu^2} \doteq 1,$$

agreeing with (20). Hence it follows that—

- (a) when  $p^2 + \nu^2 \gg l^2 p_H^2$ , the two waves are equally attenuated; and
- (b) when  $p^2 + \nu^2 \ll l^2 p_H^2$ , the extraordinary ray has the greater intensity.

\* Meaning that we confine our calculations of absorption to those regions where  $N$  is of the order unity. However, considerable portions of the atmospheric paths of long waves will be in such regions.

We immediately derive three working rules for long-wave propagation, namely :

- (a) for propagation approximately longitudinal to the earth's magnetic field, the ordinary and extraordinary downcoming rays are receivable with comparable intensities ;
- (b) for propagation approximating to the transverse type, the extraordinary ray reaches the ground with much the greater intensity ;
- (c) however, if the collision frequency is very great, for example, if the absorption occurs comparatively near to the earth's surface and not entirely where the wave is bent over in the Kennelly-Heaviside layer, then, whatever the direction of transmission, both waves are equally attenuated.

### 3. Interpretation of Hollingworth's Experimental Results.

Observations were made at Slough of the transmissions from Ste. Assise, near Paris. Approximate values of the experimental constants are given below.

Assuming the height of the Kennelly-Heaviside layer to be 80 kilometres, the ground wave path being 396 kilometres, then  $i = 68^\circ$  ; also  $\delta = 67^\circ$  and  $\gamma = 160^\circ$ , whence from (7) we have that  $\theta$ , the angle between the directions of propagation and the earth's field is  $47^\circ$ , and the propagation of the downcoming wave tends to the transverse type.

The wave-length of the Ste. Assise transmissions being 14,350 metres,  $q = 63$ , and hence the solution of (1) is :—

$R_o = 1/49$  for the ordinary ray, and

$R'_o = -49$  for the extraordinary ray.

From (6),  $S = 169^\circ$ , hence we have from (3), (4), and (5) that

$R_o = 0.19$ ,  $\xi = 6^\circ$ , ordinary, and

$R'_o = 5.2$ ,  $\xi' = 186^\circ$ , extraordinary.

Thus  $E_1$  and  $E'_1$  are very nearly in phase for the ordinary ray and in anti-phase for the extraordinary ; also the major contribution to the normal electric component  $E_1$  in the combined wave is from the ordinary ray.

We next seek information from the previous section on the differential absorption of the two rays. Since transmission approximates to the transverse type, the extraordinary ray has much the greater intensity during the night. During the day, however, when absorption can occur in the lower

atmosphere, both waves may be equally attenuated. Hence, during the sunrise period we should be prepared to use values of  $A$  between 0 and 1, the former corresponding to night conditions.

Substituting these values of  $A$ , and the known magnitude of  $R_0$  in (16), we obtain the following values for the amount of rotation of the plane of polarization :—

$$\eta = 101^\circ \text{ for night conditions, and}$$

$$\eta = 56^\circ \text{ for the steady post-sunrise value.}$$

Hollingworth's experimental results for a typical day (fig. 2 of this paper) give about  $110^\circ$  for the steady night value of  $\eta$  and  $50^\circ$  and  $60^\circ$  for the post-sunrise period. There can be very little doubt, therefore, that the analysis developed here is adequate\* to explain this rotation of the plane of polarization.

It is interesting, moreover, to see what additional information may be available. Expression (20) tells us that, during the night when the ordinary ray suffers the greater attenuation,  $p^2 + v^2 \ll l^2 p_{11}^2$ , so that the collision frequency must be much less than  $7 \times 10^6$ . Reference to Pedersen† suggests that this collision frequency might hold at a height of about 70 kilometres, so that at night-time the absorption would chiefly occur at heights greater than 70 kilometres. Now Hollingworth states that the probable height which long waves attain in the Kennelly-Heaviside layer is 75–85 kilometres. We infer that differential absorption takes place, during the night, in the region where the wave is bent over.

During the day, equality in amplitude of the two downcoming rays probably means that differential absorption occurs in a region where the collision frequency is much greater than  $7 \times 10^6$ , i.e., at heights much less than 70 kilometres. Now Hollingworth has observed that the commencement of the rapid change-over from night to day conditions is at 45 minutes before sunrise in October and 55 minutes in June, both times corresponding to a position of the sun  $7^\circ$  below the horizon for the mid-point of the path between sender and receiver. That is to say, the daytime type of differential absorption begins at the moment when it is sunrise at a height of 47 kilometres above the earth's surface. At this height the collision frequency should be, according to Pedersen, a little greater than  $10^8$  and hence much greater than our critical value of

\* We are aware of two other possible explanations of Hollingworth's results. These are discussed and rejected in the next section.

† "Propagation of Radio Waves," Copenhagen (1927), p. 43, fig. 1V, 6.

$7 \times 10^6$ . It is to be inferred that differential absorption takes place in the post-sunrise period in a region between 47 and 70 kilometres above the earth's surface.

#### 4. Discussion.

The results of the previous section suggest that the magneto-ionic theory is capable of providing an explanation of the phenomenon of the rotation of the plane of polarization of long waves, in quantitative agreement with Hollingworth's experimental observations. However, we are aware of two other interpretations of the magneto-ionic theory which could, under certain conditions, allow of a rotation of the plane of polarization. These alternatives arise when expression (1) for the limiting polarizations of the downcoming rays is modified to include collisions between electrons and other particles in the ionosphere, this involving the substitution of  $q'$  for  $q$ .

Recently, Ratcliffe\* has suggested that the effect of such a modification is that all waves, of whatever frequency, are circularly polarized at ground level, the only exception being for propagation exactly transverse to the earth's magnetic field. This interpretation of the magneto-ionic theory is in conflict with the conclusion reached by Baker and Green that electronic collisions have no very great effect on the limiting shapes of the polarization ellipses. For the purposes of the present paper, however, we would point out that Hollingworth's experiments cannot satisfactorily be aligned with Ratcliffe's interpretation of the magneto-ionic theory, reserving the discussion of the more fundamental question for another place.

An attempt to construct, on these lines, an explanation of the rotation of the plane of polarization of the combined downcoming wave, would appear to involve the assumptions that—

- (a) the amplitudes of the circularly polarized ordinary and extraordinary downcoming rays are equal, both night and day ;
- (b) the lengths of their atmospheric paths are equal during the day ;
- (c) the ordinary component lags behind the extraordinary at night.

Now we have already noted that a differential change in path-length between the two component rays is probably of second order of smallness. It is also possible to show that, for long-wave propagation, such a path difference is more likely to involve a retardation of the extraordinary ray behind the ordinary than *vice versa* as required by (c) above. Hence this alternative

\* "Wir. Eng. and Exp. Wir." vol. 10, p. 354 (1933).



explanation of Hollingworth's results seems less probable than the one we have suggested in previous sections.

The other possibility arises out of the suggestion made by Baker and Green that electronic collisions have the effect of rotating the axes of the polarization ellipse.\* Remembering that both the ordinary and the extraordinary components in the downcoming wave are nearly plane polarized, and considering each separately, it can be shown that  $\eta_{\text{ord.}}$  is  $11^\circ$  and  $\eta_{\text{ext.}}$  is  $101^\circ$  for the Ste. Assise-Slough conditions. Now the effect of collisions is to rotate the axes of each ellipse in the sense that  $\eta_{\text{ord.}}$  tends to decrease in magnitude, and  $\eta_{\text{ext.}}$  to increase, so that in neither case can we attain the full range of variation in the rotation of the plane of polarization from  $50^\circ$  to  $110^\circ$  such as was measured by Hollingworth.

At this point it will be useful to examine, from the standpoint adopted in this paper, those other long-wave anomalies which are associated with asymmetrical propagation. However, it is to be emphasized that the explanations which will be put forward must inevitably be qualitative, the lack of precision being due to the fact that the experimental observations were confined to the normal component of electric force at the receiver.

At short distances there are Naismith's observations of the Northolt transmissions in which he found that, during the daytime, the intensity of the space wave at Manchester was approximately twice as great as at Exeter. Both receiving points were distant about 240 kilometres from the transmitter, the magnetic bearings of Northolt from each of them being  $164^\circ$  and  $81^\circ$  respectively. Calculations then give  $35^\circ$  and  $63^\circ$  for the angles between the directions of propagation of the downcoming waves and the lines of force of the earth's field, so that Manchester would correspond roughly to longitudinal transmission and Exeter to transverse. It is important to remember, however, that the measurements were made in the middle of the day when there should be no differential absorption in either type of propagation. In this case we have from (11) that

$$\frac{(E_1)_M}{(E_1)_E} = \frac{1 - (R_e)_M}{1 - (R_e)_E} = \frac{1 - 0.20}{1 - 0.53} = 1.7,$$

showing a marked superiority at Manchester in the strength of the normal component of electric force, which was that observed.

Other observations made by Naismith are, at first sight, more difficult to explain in terms of the magneto-ionic theory. He found that the characteristic

\* The possibility of this effect occurring has also been noted recently by Mary Taylor, *Proc. Phys. Soc.*, vol. 45, p. 245 (1933).

sunset cycles of intensity in  $E_1$ , which are closely correlated with Hollingworth's rotation of the plane of polarization, were absent both at Manchester and at Exeter, though clearly noticeable at Aberdeen, which is geographically very much in the same direction from Northolt as is Manchester. In this case we are dealing with night conditions for which, in general, there should be differential absorption.

At Exeter, with transverse propagation, the sunset period should have been accompanied by an increase in intensity of the extraordinary ray and a rotation of the plane of polarization. However, at all times the contribution to the normal component of electric force by the extraordinary ray was small, owing partly to the value of  $R'_0$  and partly to the excess of ground wave, so that its increased intensity at sunset would have relatively little effect.

At Manchester, although the intensities of both downcoming rays were sufficiently great to make a characteristic sunset cycle observable if it occurred, yet in this case the type of propagation was nearly longitudinal, for which there is little differential absorption and consequently no very obvious sunset cycle.

We are therefore left with the observation of the sunset cycle at Aberdeen, very nearly due north of Northolt. The ground wave-path was 632 kilometres, so that the angle between the earth's field and the direction of propagation of the downcoming wave was  $53^\circ$ . Hence, although transmission was in the magnetic meridian, it was approximately of the transverse type, so that differential absorption explains the observation of the sunset cycle at Aberdeen.

In long distance propagation too little is known of the geometry of the ray paths to permit of a unique explanation of various cases. However, we may take as an example the experience of Round, Eckersley, Tremellen, and Lunnon when they were taking observations between Wyndham and Colombo of the difficulty of receiving signals from Rugby, the great circle between sender and receiver passing over the polar regions. If this single observation is typical of normal ionization conditions,\* an explanation may be suggested. It is to be supposed that propagation over such distances would be effected by a few reflections at nearly grazing incidence, for which conditions the rays would be travelling nearly normally to the earth's magnetic field for considerable distances in the polar regions. In this case  $R_0 = 0$  and  $R'_0 = \infty$ , and the normal component of electric force would be due only to the ordinary ray, which is highly attenuated. On the other hand, after combining incident and reflected waves at the receiver, the abnormal component of electric force

\* *Vide* Appleton, Naismith, and Builder, 'Nature,' vol. 132, p. 340 (1933).

is also small, owing partly to good reflection at the long wave-length but chiefly to the large angle of incidence of the downcoming rays.

The qualitative agreement found here between the experimental observations of asymmetrical long-wave propagation and the magneto-ionic theory lends support to the applicability of this hypothesis to Hollingworth's results. The agreement does not seem to be fortuitous and suggests that the procedure we have used is applicable to very long waves, although it is not obvious that methods based on a ray treatment should be adequate in such cases.

#### 5. *Acknowledgments.*

The work described here was a subsidiary investigation in the programme of the Radio Research Board of the Australian Commonwealth Council for Scientific and Industrial Research to which the authors are indebted for permission to publish their results.

Thanks are due to Dr. W. G. Baker for checking the extensions to the Baker and Green analysis, to Dr. D. F. Martyn for suggestions as to the presentation of the paper, and to the Chairman of the Radio Research Board, Professor J. P. V. Madsen, of the P.N. Russell School of Engineering of the University of Sydney, for his continued interest in the progress of the investigations.

#### 6. *Summary.*

Although the main features of long-wave propagation are well known, there have remained a number of apparent anomalies, chiefly in connection with short distance transmission, which have not yet been elucidated. Of these, the most prominent is the discovery by Hollingworth in 1927 of a regularly occurring cycle during the sunrise and sunset periods in which the plane of polarization of the downcoming wave changes rapidly from the steady night value of about  $110^\circ$  to the steady post-sunrise value of  $50^\circ$  to  $60^\circ$ .

The present paper shows that the technique of the magneto-ionic theory, hitherto used exclusively in connection with the propagation of short and medium waves, gives a feasible explanation of some long-wave anomalies, in remarkable quantitative agreement with Hollingworth's exact measurements. This agreement between theory and experiment does not seem to be fortuitous and suggests that the procedure recently advocated by Baker and Green for medium-wave propagation is also applicable to very long waves, although it is not at all obvious that methods based on a ray treatment should be adequate in such cases.

---

## *On the Electron Theory of Metals.*

By S. SCHUBIN and S. WONSOWSKY.

Sverdlovsk Physical Technical Institute.

(Communicated by R. H. Fowler, F.R.S.—Received December 29, 1933.)

### *Introduction and Summary.*

The most serious defect of the modern theory of metals† is the very unsatisfactory manner in which it takes account of the forces of interaction between the valency electrons of the metal. This is well shown in the anomalies characteristic of the electrical properties of the *ferromagnetics*; from the standpoint of Bloch's theory they are quite inexplicable, as the criterion of ferromagnetism itself can be formulated only in the language of a more accurate theory which takes account of the exchange effects.

The problem of constructing such a systematic theory of metals, which could enable us to treat their electric and magnetic properties simultaneously, reduces itself substantially to choosing a suitable approximation scheme. At first sight it would seem most natural to use the scheme applied so successfully by Heisenberg in explaining the phenomena of ferromagnetism. Here the metal is considered, in the zero approximation, as an assembly of isolated *electrically neutral* atoms; in the following approximations, account is taken of the interaction of the valency electrons not only with the ions of the metal but also with each other. In his Leipzig Report,‡ Bloch asserts that such a scheme affords an adequate tool for dealing with *all* the characteristic properties of metals, in particular with the electrical conduction. We are, however, of the opinion that this assertion is *incorrect*, and that, in reality, Heisenberg's approximation cannot be used in the theory of electrical conduction simply because, in this approximation, the metal is not a conductor. In fact, it can be proved in quite a general manner (for special cases it has already been proved by Slater§), that in all the stationary states of Heisenberg's scheme the total current carried by the valency electrons of the metal is equal to zero; this result does not depend on any one special property of the perturba-

† Bloch, 1928.

‡ 'Leipziger Vorträge,' P. Debye, Leipzig und Berlin (1930).

§ 'Phys. Rev.,' vol. 35, p. 509 (1930).

tion equations, but simply follows from the fact that only the non-polar states of the system are chosen as zero approximation, i.e., only those states in which each atom of the metal is electrically neutral. Such a theory considers only those transitions of electrons from one atom to another, in which two electrons simultaneously change places (i.e., only the exchange processes); such transitions cannot, of course, give rise to an electric current.

It is easy, however, to indicate a generalization of Heisenberg's scheme, which makes it free from the defect pointed out and permits us to obtain a method for treating the electric and magnetic properties of metals simultaneously. The idea of this generalization—which was first formulated by Slater, but has not been worked out in detail by anyone—is that, in the zero approximation, one must consider, besides the non-polar states, also the *polar* states of the system, i.e., those states where some metallic atoms have an extra electron and at the same time an equal number have lost their valency electrons. In such a scheme, besides the electrons bound to their atoms, we have a certain number of “travelling” electrons; the transitions of the latter from one atom to another—which are *not* exchange transitions, as only *one* electron takes part in each of them—make possible the propagation of the electric current. The aim of this paper is to make a systematic investigation of this polar scheme for the simplest case of a monovalent metal.

Let us consider a system of polar states with a given number of “doubly occupied” atoms  $s$ , which is also the number of “holes.” [We suppose that in all “doubly occupied” atoms both electrons are in the ground state; the class of states defined in such a manner corresponds, obviously, to the “first zone” of the usual theory.] Apparently, owing to the electrostatic forces of repulsion between the electrons, the *average* energy of the system must always increase with  $s$ . (We note that the greatest possible value of  $s$  is equal to  $n = \frac{1}{2}N$ , where  $N$  is the total number of atoms in the piece of metal considered.) The exact calculation, as we shall see, confirms this. At first sight it may seem that such a situation gives a justification for the treatment of Heisenberg, who considered only states with  $s = 0$ . In reality, however, it does not, as to each  $s$  corresponds not a single energy level, but a whole *band* of such levels and in general these bands *overlap* each other. In particular, the non-polar, “Heisenberg,” band is *always* overlapped by a certain number of polar ones. It is therefore clear that a perturbation theory which includes in the linear combinations of the zero approximation only the states with  $s = 0$  is very inexact, and in order to get better results we must *couple* the states of different  $s$  with each other in these combinations. Just this circumstance enables us, from the

standpoint of our scheme, to explain theoretically why all elements with one valency electron are metals.†

The mathematical treatment of the problem must, for the sake of convenience, be carried out in two stages. The first stage consists in the discussion of the properties of the system for a certain *fixed*  $s$ . In particular it is of interest to determine to which value of  $s$  the *lowest* energy level belongs, as, obviously, in *general*, at not very high temperatures, only those  $s$  which lie in the neighbourhood of this  $s_0$  will play an essential role. We shall see, that in many cases  $s_0 \neq 0$ , so that the lowest level belongs not to the non-polar band but to a polar one. The second stage consists in the solution of the problem for *coupled*  $s$ , which stage is of greatest interest as it corresponds best to physical reality, as the preceding argument shows. Here we must first determine the values of the energy and the current for all the stationary states of the system.

[It must be noted, that from the above standpoint, as Slater has already pointed out, the old concept of "the number of free electrons in the metal" receives a certain theoretical meaning. In fact, to each stationary state of the system corresponds a certain average value of the number  $s$ ; the statistical average of all these values can be considered as the "number of free electrons" at the given temperature. It is clear that the theory must make it possible—at least in principle—to calculate this number and determine its dependence on the temperature.]

The investigation of these general properties of the system—general character of the solutions, energy spectrum, etc.—forms the greatest part (§§ 1-3) of this first part of our paper. We shall see that, by employing methods analogous to those of Bloch,‡ one can overcome the mathematical difficulties of the problem—at least for the most important particular cases—and bring the theory into a form similar to Debye's theory of specific heat.

Another paper, now being prepared, will contain the *application* of the general results obtained to the theory of the electric and magnetic properties of metals. Some physical consequences of these general results can, however, be stated immediately here. It appears that (1) the usual criterion for ferromagnetism is *incorrect* (even for strongly bound electrons) and must be modified in the sense of a *further restriction* of the class of ferromagnetics, and (2) the elementary magnetic moment of a ferromagnetic—calculated from the value of the saturation magnetization—need *not*, in general, be a multiple of the Bohr magneton

† The question of the existence of *insulators* in this scheme will be discussed in detail elsewhere. It can be shown that the well-known arguments of Wilson can also be translated easily into the language of a many-electron theory.

‡ 'Z. Physik,' vol. 74, p. 295 (1932).

(which is just in accordance with the observed facts). Both these results are discussed in § 4.

The full report of the calculations, upon which are based the results formulated below, will be given in the 'Physikalische Zeitschrift der Sowjetunion.'

### 1. The General Equations of the Problem for Fixed $s$ .

(a) *The Perturbation Equations.*—The unperturbed wave function for a polar state is defined by specifying the "doubly occupied" atoms  $f_1, f_2, \dots, f_s$ , the "holes"  $g_1, g_2, \dots, g_s$  and finally the atoms with one right spin  $h_1, h_2, \dots, h_t$  (the atoms with one left spin  $k_1, k_2, \dots, k_u$  are then fixed automatically). Between the numbers  $s, t$ , and  $u$  exists the obvious relation

$$2s + t + u = N. \quad (1)$$

Further we have

$$s + t = \tau, \quad s + u = N - \tau, \quad (2)$$

where  $\tau$  is the total number of right spins in the system (as each double atom has one right spin and one left spin). The component of the total magnetic moment of the system along  $Ox$  is defined by the equation

$$M = (t - u) \mu = (2\tau - N) \mu = 2m\mu, \quad (3)$$

where  $\mu$  is the Bohr magneton and  $m$ , as usual, is given by

$$m = \tau - n, \quad n = \frac{1}{2}N. \quad (4)$$

The formulæ (4) and (2) show that, for a given  $s$ , the quantity  $m$  can vary only between the limits defined by the inequality

$$-(n - s) \leq m \leq n - s. \quad (5)$$

Let us denote the unperturbed wave function of the  $i$ th isolated atom of the metal (neglecting spin forces) by

$$\phi_i(x) C_\alpha(\sigma),$$

where the index  $\alpha$  can take two values R and L according to the direction of the spin  $\sigma$ . The unperturbed wave-function of the system is then the anti-symmetric combination

$$\psi_{f_1 \dots f_s, g_1 \dots g_s, h_1 \dots h_t} = \frac{1}{\sqrt{N!}} \sum_P P \phi_{f_1}(x_{f_1}) \phi_{f_2}(x_{f_2}) \dots \left. \begin{array}{l} \phi_{f_s}(x_{f_s}) \phi_{g_1}(x_{g_1}) \phi_{h_1}(x_{h_1}) \dots \phi_{k_u}(x_{k_u}) \\ C_R(\sigma_{f_1}) C_L(\sigma_{g_1}) \dots C_R(\sigma_{f_s}) C_L(\sigma_{g_s}) C_R(\sigma_{h_1}) \dots \\ C_R(\sigma_{k_u}) C_L(\sigma_{k_1}) \dots C_L(\sigma_{k_u}), \end{array} \right\} \quad (6)$$

where  $P$  denotes, as usual, the operation of permuting the co-ordinates  $x$  and

$\sigma$ ;  $\eta_P = \pm 1$  according to whether  $P$  is an even or odd permutation. As zero approximation of the solution of the Schrödinger equation

$$(\mathbf{H} - E) \psi = 0 \quad (7)$$

we must take the function

$$\psi = \sum_{fgh} C(fgh) \psi_{fgh}. \quad (8)$$

For the present we allow in these combinations only those  $(fgh)$  which correspond to *one* fixed  $s$ . The generalization to coupled  $s$  will be given in § 3.

The Hamiltonian  $H$  is given by the expression

$$H = -\frac{\hbar^2}{8\pi^2m} \sum_{i=1}^N \Delta_i + \sum_{\alpha < \beta} F_{\alpha\beta} + \sum_{\alpha, i} G_{\alpha}(x_i) + \sum_{i < k} \frac{e^2}{|x_i - x_k|}, \quad (7')$$

where  $F_{\alpha\beta}$  denotes the energy of interaction of two ions  $\alpha$  and  $\beta$ , and  $G_{\alpha}(x_i)$  the energy of electron  $i$  in the field of ion  $\alpha$ . (Here, as throughout the paper, we use the one-dimensional form of writing; the results so obtained are valid for a simple cubic lattice.)

By substituting the expression (8) into equation (7), we obtain a system of linear homogeneous equations to determine the coefficients  $C(f, g, h)$  and the energy  $E$ . In these equations appear a series of integrals, containing the atomic wave-functions  $\phi$ . For these integrals, calculated on the (inexact) supposition that different  $\phi$ 's are orthogonal to each other, we shall introduce the following notations:

$$\int \frac{e^2}{|x - x'|} \phi_{\alpha}^2(x) \phi_{\alpha}^2(x') dx dx' = A,$$

the Coulomb interaction energy between two electrons situated on *one* and the *same* atom;

$$\int \frac{e^2}{|x - x'|} \phi_{\alpha}^2(x) \phi_{\beta}^2(x') dx dx' = B_{\alpha\beta},$$

the Coulomb interaction energy between two electrons situated on *two different* atoms  $\alpha$  and  $\beta$ ;

$$\int \sum_{\gamma \neq \beta} \left[ G_{\gamma}(x) \phi_{\alpha}^2(x') + \frac{e^2}{|x - x'|} \phi_{\gamma}^2(x') \right] \phi_{\alpha}(x) \phi_{\beta}(x) dx dx' = L_{\alpha\beta},$$

the energy corresponding to the transition of an electron from atom  $\alpha$  to atom  $\beta$ ; in contrast to the exchange integral this quantity can be named a "*transition integral*" between the atoms  $\alpha$  and  $\beta$ †;

$$\int \frac{e^2}{|x - x'|} \phi_{\alpha}(x) \phi_{\beta}(x) \phi_{\alpha}(x') \phi_{\beta}(x') dx dx' = J_{\alpha\beta},$$

the usual exchange integral between the atoms  $\alpha$  and  $\beta$

† If we neglect the interaction member  $\frac{e^2}{|x - x'|}$  this integral coincides with the integral  $\beta$  of Bloch's old one-electron theory.



Besides these four types of integrals there are many other types ; for instance, the integral

$$J_{\alpha\beta\gamma} = \int \frac{e^2}{|x - x'|} \phi_{\alpha}(x) \phi_{\beta}(x) \phi_{\alpha}(x') \phi_{\gamma}(x') dx dx'.$$

This quantity, however, is essentially *smaller* than  $J_{\alpha\beta}$ , as the quantities  $\phi_{\alpha}(x) \phi_{\beta}(x)$  and  $\phi_{\alpha}(x') \phi_{\gamma}(x')$  are markedly distinct from zero in *different* regions of the lattice ; we shall therefore neglect it. We shall also neglect all integrals which contain more than *two* different pairs of functions  $\phi$ .

With these approximations (which are quite analogous to those made in the usual Heisenberg-Bloch theory), the secular equations for the coefficients  $C(fgh)$  take the form

$$\begin{aligned} \{\epsilon - s(A + D) - [\sum_{f < f'} (B_{ff'} - J_{ff'}) + \sum_{g < g'} (B_{gg'} - J_{gg'}) - \sum_{f, g} (B_{fg} + J_{fg})]\} C(fgh) \\ + \sum_{h, k} J_{hk} [C(T_{hk}|fgh) - C(fgh)] + \sum_{f, g} J_{fg} [C(T_{fg}|fgh) - C(fgh)] \\ + \sum_{f, p} L_{fp} C(T_{fp}|fgh) - \sum_{g, p} L_{gp} C(T_{gp}|fgh) = 0, \end{aligned} \quad (9)$$

where the index  $p$  runs over all the values  $h_1, \dots, h_t, k_1, \dots, k_w$ , and  $T_{\alpha\beta}|$  denotes the permutation of the two indices  $\alpha$  and  $\beta$ . The quantity  $D$  is the sum of all the exchange integrals related to any given atom ;  $\epsilon$  is the energy  $E$  of the system plus an additive constant :

$$\epsilon = E - NE_0 - C + \frac{1}{2}ND,$$

with  $E_0$  = the unperturbed energy of an isolated atom and  $C$  = the average energy of the Coulomb interaction between all the ions and all the electrons of the metal. The equations (9) form an immediate generalization of the well-known Slater equations ; by putting  $s = 0$  we return to the usual non-polar theory.

(b) *The Energetic Centre of Gravity.*—The system of equations (9) contains several independent systems, each of which corresponds to a certain given value of the quantity  $m$ , i.e., of the component of the total spin moment along  $Oz$ . It is easy to calculate the *average* value of the energy  $\epsilon$  for any such system. This value ("the energetic centre of gravity") turns out to be

$$\epsilon^*(s, m) = sA + (n^2 - m^2)D/2n. \quad (10)$$

This result shows clearly that, owing to the electrostatic repulsion between the electrons, the *average* energy of a system of polar states is greater than that of a non-polar system by the amount  $sA$ .

(c) *Translation of the Equations (9) into the Language of the Continuum.*—For some special cases the equations (9) can be solved exactly. Much better results are obtained, however, when, following Bloch,<sup>†</sup> we translate them into the language of the continuum.

Let us introduce the operators  $\Phi_i, \Psi_i, \phi_i, \psi_i$ , which satisfy the usual commutation relations of the Einstein-Bose statistics :

$$\Phi_i + \Phi_i^+ - \Phi_i, \Phi_i^+ = \delta_{ii'}, \quad \Phi_i \Psi_i - \Psi_i, \Phi_i = 0, \text{ etc.,}$$

and have the following physical meaning :—

$$\begin{aligned} \Phi_i \Phi_i^+ &= 1 && \text{when the atom } i \text{ is a "double" one,} \\ &= 0 && \text{in all other cases;} \end{aligned}$$

$$\begin{aligned} \Psi_i \Psi_i^+ &= 1 && \text{when the atom } i \text{ is a "hole,"} \\ &= 0 && \text{in all other cases;} \end{aligned}$$

$$\begin{aligned} \phi_i \phi_i^+ &= 1 && \text{when the atom } i \text{ has one right spin,} \\ &= 0 && \text{in all other cases;} \end{aligned}$$

$$\begin{aligned} \psi_i \psi_i^+ &= 1 && \text{when the atom } i \text{ has one left spin,} \\ &= 0 && \text{in all other cases.} \end{aligned}$$

(It is clear that our  $\phi_i$  and  $\psi_i$  coincide with those of Bloch.) With the aid of these operators, the equations (9) can be transformed to the "standard" type  $\epsilon C^* = H C^*$  with the Hamiltonian

$$\begin{aligned} H = & s(A + D) + \frac{1}{2} \sum_{i \neq i'} (\Phi_i \Phi_i^+ - \Psi_i \Psi_i^+) (\Phi_{i'} \Phi_{i'}^+ - \Psi_{i'} \Psi_{i'}^+) B_{ii'} \\ & - \frac{1}{2} \sum_{i \neq i'} (\Phi_i \Phi_i^+ + \Psi_i \Psi_i^+) (\Phi_{i'} \Phi_{i'}^+ + \Psi_{i'} \Psi_{i'}^+) J_{ii'} \\ & + \frac{1}{2} \sum_{i \neq i'} [(\phi_i \psi_{i'} - \phi_{i'} \psi_i) (\phi_i^+ \psi_{i'}^+ - \phi_{i'}^+ \psi_i^+) \\ & \quad + (\Phi_i \Phi_{i'} - \Phi_{i'} \Psi_i) (\Phi_i^+ \Psi_{i'}^+ - \Phi_{i'}^+ \Psi_i^+)] J_{ii'} \\ & + \frac{1}{2} \sum_{i \neq i'} [(\phi_i \phi_{i'}^+ + \psi_i \psi_{i'}^+) (\Phi_{i'} \Phi_i^+ - \Psi_{i'} \Psi_i^+) \\ & \quad + (\phi_i \phi_i^+ + \psi_i \psi_i^+) (\Phi_i \Phi_i^+ - \Psi_i \Psi_i^+)] L_{ii'}, \quad (11) \end{aligned}$$

where the index  $i$  runs over all values from 1 to  $N$ . Instead of our secular equations, we can consider the equations of motion derived from this Hamiltonian, with the obvious supplementary conditions

$$\Phi_i \Phi_i^+ + \Psi_i \Psi_i^+ + \phi_i \phi_i^+ + \psi_i \psi_i^+ = 1, \quad (12)$$

$$\sum_{i=1}^N \Phi_i \Phi_i^+ = \sum_{i=1}^N \Psi_i \Psi_i^+ = s. \quad (13)$$

<sup>†</sup> 'Z. Physik,' vol. 74, p. 295 (1932).

To make these formulæ of practical use we have to make two fundamental simplifications, similar to those made by Bloch in § 4 of his last cited paper. The limits of their validity will be discussed later. The simplifications are as follows: (1) in the sums of expression (11) we neglect all members except those with  $i' = i + 1$ ; (2) instead of the discrete set of operators  $\Phi_i, \Psi_i, \phi_i, \psi_i$ , we introduce four continuous functions  $\Phi(x), \Psi(x), \phi(x), \psi(x)$ , by putting

$$\Phi_i = \Phi(x), \quad \Phi_{i \pm 1} = \Phi(x) \pm a \frac{d\Phi}{dx} + \frac{a^2}{2} \frac{d^2\Phi}{dx^2},$$

where  $a$  is the lattice constant; these functions we consider as simple  $c$ -numbers, i.e., we neglect the commutation relations between  $\Phi, \Phi^\dagger$ , etc.

Using these simplifications and passing from sums to integrals, we immediately obtain

$$\begin{aligned} H = & s(A + 6J) + \frac{B}{a} \left[ \frac{1}{a^2} \int (\Phi \Phi^* - \Psi \Psi^*)^2 dx - \frac{1}{2} \int (\Phi \Phi^* - \Psi \Psi^*)^2_a dx \right] \\ & - \frac{J}{a} \left[ \frac{1}{a^2} \int (\Phi \Phi^* + \Psi \Psi^*)^2 dx - \frac{1}{2} \int (\Phi \Phi^* + \Psi \Psi^*)^2_a dx \right] \\ & + \frac{J}{a} \int \left[ |\phi \phi_a - \psi \psi_a|^2 + |\Phi \Psi_a - \Psi \Phi_a|^2 \right] dx \\ & + \frac{2L}{a} \left[ \frac{1}{a^2} \int (\phi \phi^* + \psi \psi^*) (\Phi \Phi^* - \Psi \Psi^*) dx \right. \\ & \left. - \frac{1}{2} \int (|\langle \phi \Phi^* \rangle_a|^2 + |\langle \psi \Phi^* \rangle_a|^2 - |\langle \phi \Psi^* \rangle_a|^2 - |\langle \psi \Psi^* \rangle_a|^2) dx, \right] \quad (14) \end{aligned}$$

and

$$\Phi \Phi^* + \Psi \Psi^* + \phi \phi^* + \psi \psi^* = 1, \quad (15)$$

$$\frac{1}{a^2} \int \Phi \Phi^* dx = \frac{1}{a^2} \int \Psi \Psi^* dx = s, \quad (16)$$

where the suffix  $x$  denotes the derivative  $d/dx$  and where we have put

$$B_{i, i+1} \equiv B_{i, i-1} = B, \quad J_{i, i+1} \equiv J_{i, i-1} = J, \quad L_{i, i+1} = L_{i, i-1} = L.$$

For the stationary states of the system, i.e., when the functions  $\Phi, \Psi, \phi, \psi$ , have the form

$$\begin{aligned} \Phi(x, t) = U(x) e^{2\pi i \lambda t / \hbar}, \quad \Psi(x, t) = V(x) e^{2\pi i \lambda x / \hbar}, \quad \phi(x, t) = u(x) e^{2\pi i \lambda x / \hbar}, \\ \psi(x, t) = v(x) e^{2\pi i \lambda x / \hbar} \end{aligned}$$

the equations of motion derived from expression (14) can be written as follows :

$$\begin{aligned}
 & 2B\Phi \left[ (\Phi\Phi^* - \Psi\Psi^*) + \frac{a^2}{2}(\Phi\Phi^* - \Psi\Psi^*)_{xx} \right] \\
 & - 2J\Phi \left[ (\Phi\Phi^* + \Psi\Psi^*) + \frac{a^2}{2}(\Phi\Phi^* + \Psi\Psi^*)_{xx} \right] \\
 & + Ja^2[2(\Phi\Psi_x - \Psi\Phi_x)\Psi^* + (\Phi\Psi_{xx} - \Psi\Phi_{xx})\Psi^*] \\
 & - 2L \left\{ \phi \left[ \Phi\phi^* + \frac{a^2}{2}(\Phi\phi^*)_{xx} \right] + \psi \left[ \Phi\psi^* + \frac{a^2}{2}(\Phi\psi^*)_{xx} \right] \right\} \\
 & \hspace{15em} = -\lambda_1\Phi \\
 & 2B\Psi \left[ (\Phi\Phi^* - \Psi\Psi^*) + \frac{a^2}{2}(\Phi\Phi^* - \Psi\Psi^*)_{xx} \right] \\
 & + 2J\Psi \left[ (\Phi\Phi^* + \Psi\Psi^*) + \frac{a^2}{2}(\Phi\Phi^* + \Psi\Psi^*)_{xx} \right] \\
 & + Ja^2[2(\Phi\Psi_x - \Psi\Phi_x)\Phi^* + (\Phi\Psi_{xx} - \Psi\Phi_{xx})\Phi^*] \\
 & - 2L \left\{ \phi \left[ \Psi\phi^* + \frac{a^2}{2}(\Psi\phi^*)_{xx} \right] + \psi \left[ \Psi\psi^* + \frac{a^2}{2}(\Psi\psi^*)_{xx} \right] \right\} \\
 & \hspace{15em} = \lambda_2\Psi \\
 & Ja^2[2(\phi\psi_x + \psi\phi_x)\psi^* + (\phi\psi_{xx} - \psi\phi_{xx})\psi^*] \\
 & - 2L \left\{ \Phi \left[ \phi\Phi^* + \frac{a^2}{2}(\phi\Phi^*)_{xx} \right] - \Psi \left[ \phi\Psi^* + \frac{a^2}{2}(\phi\Psi^*)_{xx} \right] \right\} \\
 & \hspace{15em} = -\lambda_3\phi \\
 & Ja^2[2(\phi\psi_x - \psi\phi_x)\phi^* + (\phi\psi_{xx} - \psi\phi_{xx})\phi^*] \\
 & + 2L \left\{ \Phi \left[ \psi\Phi^* + \frac{a^2}{2}(\psi\Phi^*)_{xx} \right] - \Psi \left[ \psi\Psi^* + \frac{a^2}{2}(\psi\Psi^*)_{xx} \right] \right\} \\
 & \hspace{15em} = \lambda_4\psi.
 \end{aligned} \tag{17}$$

It is easy to determine the connection between the parameters  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ , and the energy of the system  $H$ . In fact, if we multiply the first of the equations (17) by  $\Phi^*$ , the second by  $-\Psi^*$ , the third by  $\phi^*$ , the fourth by  $-\psi^*$ , add and integrate over  $dx/a^3$ , we easily obtain

$$H = s(A + 6J) - \left( s \frac{\lambda_1 + \lambda_2}{2} + (n - s) \frac{\lambda_2 + \lambda_4}{2} + m \frac{\lambda_3 - \lambda_4}{2} \right), \tag{18}$$

as

$$\begin{aligned}
 & \frac{1}{a^3} \int (\Phi\Phi^* + \Psi\Psi^*) dx = 2s, \quad \frac{1}{a^3} \int (\Phi\Phi^* - \Psi\Psi^*) dx = 0 \\
 & \frac{1}{a^3} \int (\phi\phi^* + \psi\psi^*) dx = 2(n - s), \quad \frac{1}{a^3} \int (\phi\phi^* - \psi\psi^*) dx = 2m
 \end{aligned} \tag{19}$$

The formulæ (14)–(18) will serve us as a basis for the discussion of the most important special cases of the problem.

## 2. The Solution of Equations (17) for some Special Cases.

(a) *The case  $s = 0$ .*—In order to illustrate the general method of solution of equations (17), we shall discuss here in some detail the non-polar case  $s = 0$ . For this we have  $\Phi = \Psi = 0$ , and our formulæ (14), (15), and (17) become those of Bloch :

$$H = \frac{J}{a} \int |\phi\psi_x - \psi\phi_x|^2 dx, \quad (20)$$

$$\phi\phi^* + \psi\psi^* = 1, \quad (21)$$

$$\left. \begin{aligned} Ja^2 [2(\phi\psi_x - \psi\phi_x)\psi_x^* + (\phi\psi_{xx} - \psi\phi_{xx})\psi^*] &= -\lambda_3\phi \\ Ja^2 [2(\phi\psi_x - \psi\phi_x)\phi_x^* + (\phi\psi_{xx} - \psi\phi_{xx})\phi^*] &= \lambda_4\psi \end{aligned} \right\}, \quad (22)$$

furthermore we have

$$H = -\frac{1}{2} [(\lambda_3 + \lambda_4)n + (\lambda_3 - \lambda_4)m]. \quad (23)$$

If we multiply the first of the equations (22) by  $\phi^*$ , the second by  $\psi^*$  and integrate by parts, we obtain

$$\begin{aligned} Ja^2 \int |\phi\psi_x - \psi\phi_x|^2 dx &= -\lambda_3 \int \phi\phi^* dx, \\ -Ja^2 \int |\phi\psi_x - \psi\phi_x|^2 dx &= \lambda_4 \int \psi\psi^* dx, \end{aligned}$$

or

$$\lambda_3 \int \phi\phi^* dx = \lambda_4 \int \psi\psi^* dx, \quad \lambda_3 - \lambda_4 = -m(\lambda_3 + \lambda_4)/n. \quad (24)$$

The expression (23) then becomes

$$H = -(\lambda_3 + \lambda_4) \cdot (n^2 - m^2)/2n. \quad (25)$$

To calculate  $\lambda_3 + \lambda_4$  we must use the condition (21), which enables us to seek the solution of equations (22) in the form

$$\phi = \sin y e^{i\alpha}, \quad \psi = \cos y e^{i\beta}, \quad (26)$$

The substitution of this in (22) leads to the equations

$$\left. \begin{aligned} Ja^2 (2 \sin y \cdot y'^2 - \cos y \cdot y'' + \sin y \cos^2 y \cdot \gamma'^2) &= -\lambda_3 \sin y \\ Ja^2 (2 \cos y \cdot y'^2 + \sin y \cdot y'' + \sin^2 y \cos y \cdot \gamma'^2) &= -\lambda_4 \cos y \\ \frac{1}{2} \sin 2y \cdot \gamma'' + 2 \cos 2y \cdot y' \gamma' &= 0 \end{aligned} \right\}, \quad (27)$$

where  $\gamma = \beta - \alpha$ . Multiplying the first of equations (27) by  $\sin y$ , the second by  $\cos y$  and adding, we obtain

$$-Ja^2(2\cos 2y \cdot y'^2 + \sin 2y \cdot y'') = -\frac{1}{2}(\lambda_3 - \lambda_4) + \frac{1}{2}(\lambda_3 + \lambda_4)\cos 2y$$

or

$$J \frac{d^2}{dx^2}(\cos 2y) = (\lambda_3 + \lambda_4)(\cos 2y + m/n).$$

This equation can be solved immediately ;

$$\cos 2y = -m/n + p \cos(\omega x + q) \quad (28)$$

with

$$\lambda_3 + \lambda_4 = -J\omega^2 a^2, \quad (29)$$

and for the energy of the system we obtain

$$H = (n^2 - m^2) J \omega^2 a^2 / 2n. \quad (30)$$

The formulæ (28)–(30) are valid for a linear chain of atoms. For a three-dimensional cubic lattice we have instead

$$\cos 2y = -m/n + p \cos(\omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 + q) \quad (31)$$

and

$$H = (n^2 - m^2) J (\omega_1^2 + \omega_2^2 + \omega_3^2) a^2 / 2n. \quad (32)$$

The consequences of expressions (31) and (32) allow us to establish very easily the validity of the approximation employed. In fact, the quantities  $\omega_1$ ,  $\omega_2$ ,  $\omega_3$  are the frequencies of the spin waves  $\phi\phi^*$  and  $\psi\psi^*$  along the axes  $x_1$ ,  $x_2$ ,  $x_3$  (as  $\phi\phi^* = \frac{1}{2}(1 - \cos 2y)$ ,  $\psi\psi^* = \frac{1}{2}(1 + \cos 2y)$ ). From the standpoint of the "continuum theory" these frequencies can be as large as we like; therefore the energy—given by (32)—can take all possible values from 0 to  $\infty$ . This circumstance is quite analogous to the appearance of infinitely large frequencies in Debye's theory of specific heats, since the transition from the exact expression (11) to (14) is itself *quite similar to the transitions from Born's exact theory to Debye's theory*. These considerations show, that to obtain consistent results from expression (32), we must artificially cut off the spectrum of the frequencies  $\omega$  by taking into account the atomic

FIG. 1.

constitution of the metal. The maximum  $\omega$ , in our case, will obviously correspond to the "most non-uniform" distribution of spins, shown in fig. 1, giving

$$\omega_1^{(\max)} = \omega_2^{(\max)} = \omega_3^{(\max)} = \pi/a. \quad (33)$$

The formulæ (32) and (33) give a complete description of the energy spectrum of the system (to our approximation). They can be used, for example, as the basis for a theory of ferromagnetism, neglecting spin forces, valid for all temperatures and field-strengths.

Expression (32) shows that the minimum energy level corresponds to  $\omega_1 = \omega_2 = \omega_3 = 0$ —uniform distribution of spins—so we have for all  $m$

$$H_{\min} = 0. \quad (34)$$

For the maximum, however, we have

$$H_{\max}(m) = 3\pi^2(n^2 - m^2)J/2n. \quad (35)$$

These maxima are different for different  $m$ ; they have the same relationship to each other as the corresponding averages (10). [The above, obviously, is correct only if  $J > 0$ ; if  $J < 0$  we must interchange the words “maximum” and “minimum”]

It must be noted that, to make the maximum energies (35) agree with those calculated *exactly* for some special  $m$  by Bloch,<sup>†</sup> we have to replace the factor  $\pi^2$  by 4; this factor  $\pi^2/4$  appears in quite an analogous manner if one compares the results of the exact theory of Born with those of Debye. To make the formulæ more symmetrical, we shall make this substitution and write here, and in all similar cases in the future, 4 instead of  $\pi^2$ :

$$H_{\max}(m) = 6(n^2 - m^2)J/n \quad (36)$$

the change, of course, being of no importance. We have then for the energetic centre of gravity

$$\overline{\epsilon(0, m)} = \frac{1}{2} \{H_{\min}(m) + H_{\max}(m)\}.$$

Finally, we must call attention to the fact that the expression (32) for the energy has been obtained quite independently of the values which could be given to the constants  $p$  and  $q$ . The phase  $q$  has not, of course, any physical significance; as to  $p$ , it must be chosen so as to make our solution (31) satisfy *all* the equations (27). For instance, for  $m = 0$ , we must put

$$p = 1, \quad q = \text{const.}, \quad y = \frac{1}{2}(\omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 + q),$$

and equations (27) will be satisfied by the functions

$$\phi = \sin \frac{1}{2}(\omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 + q), \quad \psi = \cos \frac{1}{2}(\omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 + q). \quad (37)$$

<sup>†</sup> ‘Z. Physik,’ vol. 61, p. 206 (1930).

(b) *The case  $s = n$ .*—For this case equations (9) can be solved exactly if we neglect the exchange integrals  $J$ . To solve the problem without doing this we must employ the approximate equations (17); the method is quite similar to that of the preceding section. We obtain, finally, for  $\Phi$  and  $\Psi$ , expressions of the type (37), and the maximum and minimum energies are given by

$$H_{\min} = n(A - 6B) + 6nJ \quad (38)$$

$$H_{\max} = n(A + 6B). \quad (38')$$

If  $J > 2B$  the roles of maximum and minimum must be inverted. As a certain verification of the calculations, we note that the average energy given by (10) for this case coincides with the arithmetical mean of  $H_{\min}$  and  $H_{\max}$ . The physical meaning of these formulæ can be understood very easily; the energy (38) corresponds to the distribution in fig. 2. As compared with the

FIG. 2.

distribution of fig. 1, each atom has *two* electrons, giving an energy increase of  $nA$ —but at the same time each electron has lost its neighbours in the neighbouring atoms, with a consequent decrease in energy of  $znB$ , where  $z$  is the number of neighbouring atoms. That, however, is just the result (38) (except for the exchange correction, whose value, of course, cannot have any physical meaning in a classical picture). By a similar argument the result (38') can be made clear. This corresponds to the distribution represented in fig. 3.

FIG. 3.

(c) *The case  $s \ll n$ .*—We have not been able to find the general solution of equation (17). One can, however, obtain a very good idea about the character of these solutions by the method of successive approximations. In fact, we can put for the zero approximation  $s = 0$  or  $s = n$ , and then seek the solution in the form of a power series in  $s/n$  or  $(n - s)/n$ . We shall discuss here the results obtained on neglecting all powers of  $s/n$  except the first.

Just as in the cases (a) and (b), the solution without exchange correction can be obtained here starting from the exact equations (9). This solution is closely analogous to that considered by Bloch (*loc. cit.*). For instance, for the case  $s = 1$  we have

$$\epsilon = A - 2L(\cos \xi_1 + \cos \eta_1 + \cos \zeta_1 - \cos \xi_2 - \cos \eta_2 - \cos \zeta_2).$$



If we neglect the interaction member  $e^2/|x - x'|$  in the integrals A and L (i.e., simply strike off the integral A), this expression becomes that of Bloch (with a somewhat different notation), with only this difference, that we have  $\cos \xi_1 - \cos \xi_2$  instead of Bloch's  $\cos \xi$ . This result can be very well understood, because, in fact, the case  $s = 1$  corresponds to just the problem considered by Bloch. Our *two* sets of variables  $\xi, \eta, \zeta$  and  $\xi_2, \eta_2, \zeta_2$  arise because we consider not only an electron but also a *hole* moving in the field of ions of the metal.

The exchange integrals can be taken into account with the help of equations (17). We have carried out these calculations for two cases: (1) when the metal is unmagnetized, i.e.,  $\phi$  and  $\psi$  are given, in the zero approximation, by (37), and (2) when the spins are uniformly distributed, for arbitrary  $m$ , i.e.,  $\phi$  and  $\psi$  in the zero approximation are constant. We obtain, then, for  $\Phi$  and  $\Psi$ , solutions of the type of plane waves:

$$\Phi = ae^{i(\nu'_1 x_1 + \nu'_2 x_2 + \nu'_3 x_3)}, \quad \Psi = be^{i(\nu''_1 x_1 + \nu''_2 x_2 + \nu''_3 x_3)} \quad (39)$$

which correspond to the free propagation of electrons and holes through the metal. In higher approximations this propagation will be disturbed by the Coulomb interaction between the "free" electrons, i.e., by the members with the integral B. As to the spin waves  $\phi$  and  $\psi$ , their amplitude in the first approximation will be slightly diminished (in the ratio  $\sqrt{(n-s)/n}$ ); for some cases they can also be partly modulated with the frequencies  $2\nu'_1, 2\nu''_1$ , etc. The minimum and maximum energies are given by the following expressions (the terms "minimum" and "maximum" refer here—as also in the following section—to the case  $J > 0$ ; for  $J < 0$  they must be inverted):

$$H_{\min} = s(A + 6J - 12L) \quad (40)$$

$$H_{\max} = s(A - 6J + 12L) + 6nJ. \quad (40')$$

The level (40) belongs to all  $m$  which are possible for the given value of  $s$  (formula (5)) and the level (40') only to  $m = 0$ . It can again be verified easily that

$$\overline{\varepsilon(s, m)} = \frac{1}{2}(H_{\min} + H_{\max}).$$

(d) *Discussion.*—The most interesting point to be decided is the relative positions of the energy minima  $H_{\min}$ . This can be done by comparing formulæ (34), (38), and (40). We see, that in spite of the fact that the *average* energy of a polar system of states is always greater than that of the non-polar system, the situation can be reversed for the *lowest* levels. Formulæ (34),

(38), and (40) show that, in our scheme, three different possibilities can be realized :—

(I). The minimum energy corresponds to  $s = n$ . This is so when

$$A + 6(J - B) < 0,$$

and, for instance,

$$A + 6J - 12L > 0$$

(it is to be noted that the quantities  $A$ ,  $B$ , and  $L$  are, in general, of the same order of magnitude). For this case one can obviously expect good results from the usual theory, which puts the number of free electrons equal to the number of atoms. This possibility is probably realized only for the alkali metals whose properties are described very well by the Peierls approximation of “almost free” electrons.

(II). The minimum energy corresponds to a certain  $s = s_0$ , where  $0 < s_0 < n$ . This case we have, for instance, when

$$A + 6(J - B) > 0, \quad A + 6J - 12L < 0.$$

Then, so long as  $s$  remains small, the lowest energy level *diminishes* as  $s$  increases; for a certain  $s = s_0$  it attains a minimum and then again begins to increase. For such metals—at not very high temperatures—the number of “free” electrons approximates to twice this  $s_0$  (electrons + holes!) and is therefore *smaller* than the number of atoms. In order to calculate  $s_0$  in terms of our integrals, the energy must be evaluated up to the second approximation in powers of  $s/n$ ; we shall not, however, make these rather cumbersome calculations here.

(III). The minimum energy corresponds to  $s = 0$ . This is the case when

$$A + 6(J - B) > 0, \quad A + 6J - 12L > 0.$$

For this case, the coupling of states with different  $s$  will, in general, have more influence on the energy spectrum than for the two preceding cases. If this coupling does not essentially alter the distribution of the lowest energies (§ 3), one can say that the number of free electrons will be small and strongly dependent on the temperature.

[To avoid misunderstanding, we must insist here on the meaning which we imply, as also in other theories, by the term “number of free electrons.” This must not be confused with the number of electrons which take part in the electric current, as this number is always extremely small. When we say, for instance, in Sommerfeld’s theory or in Bloch’s old theory, that the number

of free electrons has a definite value  $s$ , we mean only that when we calculate the maximum energy of these electrons (at the absolute zero), we consider  $s/2$  lowest cells in the phase space as occupied. In our scheme the following corresponds to such a *definition*. Let us neglect the exchange forces and consider the difference between the maximum and minimum energies of the system—the breadth of the energy band. This difference can be represented as the product of a certain number  $s$  and an energy related to *one* electron, for instance, a combination of  $A$ ,  $6B$ ,  $6L$ , etc. This number is just a measure of the number of free electrons.]

In the following paragraph we shall see that, when the coupling of different  $s$  is taken into account, although the quantitative criteria formulated above are somewhat modified—the general character of these considerations is unaltered. We are, in particular, convinced that the division of metals into three classes, introduced above—especially the division between metals of class I and those of classes II and III—has a real physical significance. As shown in § 4, it throws a certain new light on the explanation of the magnetic properties of metals.

### 3. The Coupling of States of different $s$ .

(a) *General Equations.*—Let us now consider the effect of introducing into the linear combinations (8) the wave-functions corresponding to *different* values of the number  $s$ . With our approximations, this generalization will give us the following two new members on the right-hand side of equation (9) :

$$\begin{aligned} & - \sum_{h,k} L_{hk} \{C(S_{h \rightarrow f, k \rightarrow g} | fgh) - C(S_{k \rightarrow f, h \rightarrow g} | fgh)\} \\ & - \sum_{f,g} L_{fg} \{C(S_{f \rightarrow h, g \rightarrow k} | fgh) - C(S_{f \rightarrow k, g \rightarrow h} | fgh)\}, \end{aligned} \quad (9')$$

where  $S_{h \rightarrow f, k \rightarrow g}$  denotes an operator which transforms an atom  $h$  with right spin into a double atom  $f$ , and at the same time transforms an atom  $k$  with left spin into a hole  $g$ ; the other symbols  $S$  have a similar meaning. The physical meaning of the new members is that we consider now the process of a simultaneous "birth" (or "annihilation") of a double atom and a hole, in addition to the processes of "travelling" of double atoms and holes.

The members (9') give the following supplementary member to the Hamiltonian (11) :

$$\begin{aligned} & - \frac{1}{2} \sum_{i \neq i'} \{(\Phi_i \Psi_{i'} - \Phi_{i'} \Psi_i)(\phi_i^+ \psi_{i'}^+ - \psi_i^+ \phi_{i'}^+) + (\Phi_i^+ \Psi_{i'}^+ \\ & - \Phi_{i'}^+ \Psi_i^+)(\phi_i \psi_{i'} - \psi_i \phi_{i'})\} L_{ii'}, \end{aligned} \quad (11')$$

or, in the language of the continuum :

$$-\frac{L}{a} \int \{ (\Phi \Psi_s - \Psi \Phi_s) (\phi^* \psi_s^* - \psi^* \phi_s^*) + (\phi \psi_s - \psi \phi_s) (\Phi^* \Psi_s^* - \Psi^* \Phi_s^*) \} dx, \quad (14')^\dagger$$

The full equations of motion for the case of coupled  $s$  are :

$$\left. \begin{aligned} & 2B\Phi \left[ (\Phi\Phi^* - \Psi\Psi^*) + \frac{a^2}{2} (\Phi\Phi^* - \Psi\Psi^*)_{xx} \right] \\ & \quad - 2J\Phi \left[ (\Phi\Phi^* + \Psi\Psi^*) + \frac{a^2}{2} (\Phi\Phi^* + \Psi\Psi^*)_{xx} \right] \\ & \quad + Ja^2 [2(\Phi\Psi_s - \Psi\Phi_s) \Psi_s^* + (\Phi\Psi_{xx} - \Psi\Phi_{xx}) \Psi^*] \\ & \quad - 2L \left\{ \phi \left[ \Phi\phi^* + \frac{a^2}{2} (\Phi\phi^*)_{xx} \right] + \psi \left[ \Phi\psi^* + \frac{a^2}{2} (\Phi\psi^*)_{xx} \right] \right\} \\ & \quad - La^2 [2(\phi\psi_s - \psi\phi_s) \Psi_s^* + (\phi\psi_{xx} - \psi\phi_{xx}) \Psi^*] = -\lambda_1 \Phi \\ & 2B\Psi \left[ (\Phi\Phi^* - \Psi\Psi^*) + \frac{a^2}{2} (\Phi\Phi^* - \Psi\Psi^*)_{xx} \right] \\ & \quad + 2J\Psi \left[ (\Phi\Phi^* + \Psi\Psi^*) + \frac{a^2}{2} (\Phi\Phi^* + \Psi\Psi^*)_{xx} \right] \\ & \quad + Ja^2 [2(\Phi\Psi_s - \Psi\Phi_s) \Phi_s^* + (\Phi\Psi_{xx} - \Psi\Phi_{xx}) \Phi^*] \\ & \quad - 2L \left\{ \phi \left[ \Psi\phi^* + \frac{a^2}{2} (\Psi\phi^*)_{xx} \right] + \psi \left[ \Psi\psi^* + \frac{a^2}{2} (\Psi\psi^*)_{xx} \right] \right\} \\ & \quad - La^2 [2(\phi\psi_s - \psi\phi_s) \Phi_s^* + (\phi\psi_{xx} - \psi\phi_{xx}) \Phi^*] = \lambda_2 \Psi \\ & Ja^2 [2(\phi\psi_s - \psi\phi_s) \psi_s^* + (\phi\psi_{xx} - \psi\phi_{xx}) \psi^*] \\ & \quad - 2L \left\{ \Phi \left[ \phi\Phi^* + \frac{a^2}{2} (\phi\Phi^*)_{xx} \right] - \Psi \left[ \phi\Psi^* + \frac{a^2}{2} (\phi\Psi^*)_{xx} \right] \right\} \\ & \quad - La^2 [2(\Phi\Psi_s - \Psi\Phi_s) \psi_s^* + (\Phi\Psi_{xx} - \Psi\Phi_{xx}) \psi^*] = -\lambda_3 \phi \\ & Ja^2 [2(\phi\psi_s - \psi\phi_s) \phi_s^* + (\phi\psi_{xx} - \psi\phi_{xx}) \phi^*] \\ & \quad - 2L \left\{ \Phi \left[ \psi\Phi^* + \frac{a^2}{2} (\psi\Phi^*)_{xx} \right] - \Psi \left[ \psi\Psi^* + \frac{a^2}{2} (\psi\Psi^*)_{xx} \right] \right\} \\ & \quad - La^2 [2(\Phi\Psi_s - \Psi\Phi_s) \phi_s^* + (\Phi\Psi_{xx} - \Psi\Phi_{xx}) \phi^*] = \lambda_4 \psi \end{aligned} \right\} \quad (17')$$

† At first sight it may seem rather strange that the expressions (9'), (11'), and (14') are *antisymmetric* rather than symmetric relative to a permutation of right and left spins. This asymmetry has, however, only a *formal* character; it arises from our choice of the first member in the sum (6). In fact, instead of

$$\phi_{r_1}(x_{r_1}) \phi_{l_1}(x_{l_1}) C_R(\sigma_{r_1}) C_L(\sigma_{l_1}),$$

we could also write

$$\phi_{l_1}(x_{l_1}) \phi_{r_1}(x_{r_1}) C_L(\sigma_{r_1}) C_R(\sigma_{l_1}).$$

One could easily verify, that it would just change the sign in the expressions (9'), (11'), and (14'). This circumstance has no physical consequences.

where  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$  are constant parameters, connected with the energy  $H$  by the same relation (18) as before. Here, however, we must understand by  $s$  simply the quantity

$$\frac{1}{2a^3} \int (\Phi \Phi^* + \Psi \Psi^*) dx,$$

whose value, in general, can be *different* for different stationary states of the system.

[It might seem somewhat strange that, in spite of coupling, we ascribe a *definite* value of the number  $s$  to each stationary state of the system. This, however, is a logical consequence of the transition from the exact expression (11)–(11') to the “semi-classical” (14)–(14'). In fact, the coupling of different  $s$  is expressed mathematically by the non-commutability of the operator  $s$ ; defined by (13), with the Hamiltonian (14)–(14'). It is, however, just that non-commutability which we neglect in the semi-classical theory. The quantity

$$\frac{1}{2a^3} \int (\Phi \Phi^* + \Psi \Psi^*) dx$$

is therefore a certain average of all those  $s$  which enter into the corresponding linear combination (8).]

(b) *The case  $s \ll n$ .*—We shall consider here the solution of equation (17') for the case  $s \ll n$ ; the formulæ given below—like those of § 2 (c)—are exact only up to terms of order  $s/n$ .

In this approximation the functions  $\Phi$  and  $\Psi$  will have the form

$$\Phi = a_1 e^{i\nu_1 x} + a_2 e^{i\nu_2 x}, \quad \Psi = b_1 e^{-i\nu_1 x} + b_2 e^{-i\nu_2 x}, \quad (41)$$

i.e., will be more complicated than for a small *fixed*  $s$  (formulæ (39)). The coefficients  $a_1, a_2$  are connected with the corresponding  $b_1, b_2$  by definite linear relations; so that, in general, we cannot have here, for instance,  $\Psi = 0$  when  $\Phi \neq 0$ . This can easily be understood physically: there will always exist, in fact, a certain finite probability of a simultaneous birth of a double atom and a hole.

The spin waves  $\phi$  and  $\psi$  will have a form here similar to that of § 2 (c); their amplitude will be diminished in the same ratio  $\sqrt{(n-s)/n}$ , and further, they will be partly modulated by the frequencies  $\nu_1 - \nu_2$ ,  $2\nu_1$ ,  $2\nu_2$ , and  $2(\nu_1 - \nu_2)$ .

The question of the situation of the energy minimum  $H_{\min}$  affords the greatest interest here, as before. We must distinguish a great many special cases.

(I)  $J > 0$ .—The energy minimum is given by the smallest of the following three expressions

$$(I') \quad H_{\min} = 0 \quad (42')$$

belonging to  $s = 0$  and to all  $m$  from  $-n$  to  $+n$ .

$$(I'') \quad H_{\min} = s_0 (A + 6J - 12L) \quad (42'')$$

where the quantity  $s_0$  is the same as that defined in § 2 (c). This level belongs to  $s = s_0$  and to all  $m$  satisfying the inequality

$$-(n - s_0) \leq m \leq n - s_0. \quad (42_1'')$$

For the other values of  $m$   $H_{\min}$  will be greater :

$$H_{\min} = (n - |m|) (A + 6J - 12L) \quad (42_2'')$$

where  $n - |m|$  is the greatest  $s$  possible for a given  $m$  (formula (5)).

$$(I''') \quad H_{\min} = s_1 (A - 6J - 12\sqrt{5}L) + 6nJ \quad (42''')$$

belonging to a certain  $s_1 > 0$ , whose value can be determined by carrying out the calculations up to the second approximation in powers of  $s/n$ , and to  $m = 0$ . For all other  $m$ ,  $H_{\min}$  will be greater.

The cases (I') and (I'') are what would be expected from (34) and (40) of § 2. The possibility (I''') appears only on account of the coupling ; it shows that, in spite of  $J > 0$ , the lowest energy level can belong to the band  $m = 0$ .

(II)  $J < 0$ .—Here the minimum is the smaller of the two expressions :

$$(II') \quad H_{\min} = 6nJ \quad (43')$$

belonging to  $s = 0$  and  $m = 0$ ,

$$(II'') \quad H_{\min} = s_2 (A - 6J - 12\sqrt{5}L) + 6nJ \quad (43'')$$

corresponding to a certain  $s_2 > 0$  and to  $m = 0$ .

(c) *Other Special Cases.*—We shall not discuss here the case of  $s$  near to  $n$ , but note only that, if the lowest level belongs to  $s = n$ , it belongs also to  $m = 0$  as this is the one and only value of  $m$  possible when  $s = n$ .

(d) *Discussion.*—It is interesting to compare the results formulated above with those of the usual Heisenberg-Bloch theory. In the latter there were

only two possibilities, represented graphically in figs. 4 and 5 (where the abscissæ are the values of  $m$  and the ordinates the energy minima) :—

- ( $\alpha$ ) The energy minimum belongs to *all*  $m$  from  $-n$  to  $+n$ .
- ( $\beta$ ) The energy minimum belongs only to the band  $m = 0$ .

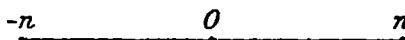


FIG. 4.

The first case occurred always for  $J > 0$ , the second for  $J < 0$ . In the polar theory, the situation is modified in two respects. Firstly, the chances of realization of case ( $\beta$ ) are now much larger; we can very well have the lowest level in the band  $m = 0$  even with a positive exchange integral. Case ( $I''$ )

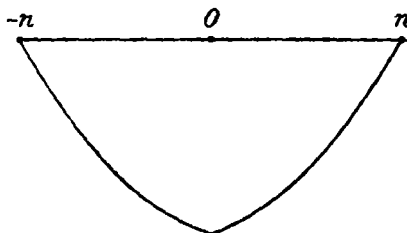


FIG. 5.

of section (b) can be cited as an example and also the case  $s_0 = n$  (number of free electrons = number of atoms!), where we always have  $m = 0$  for the lowest state, quite independently of the sign of the exchange integral. Secondly, we have here the *third* possibility ( $\gamma$ ) (case  $I''$ ), represented graphically in fig. 6, in which the lowest level corresponds not to all  $m$  but to a certain range —  $(n - s_0)$  to  $(n - s_0)$  of  $m$ . Both these corrections are of great importance in the theory of the magnetic properties of metals.

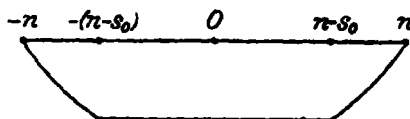


FIG. 6.

#### 4. Applications to Ferromagnetism.

According to Heisenberg a metal is ferromagnetic when, at the absolute zero of temperature, it can be magnetized up to saturation without any expense of energy; otherwise it is a paramagnetic (or a diamagnetic). In other words, for a paramagnetic the lowest energy level must correspond to the un-

magnetized state ( $m = 0$ ); for a ferromagnetic the lowest level must remain the same at least for a certain range of values of  $m$ , at least neglecting spin forces.

In the old Heisenberg-Bloch theory the first possibility was realized always for  $J < 0$ , the second always for  $J > 0$ . Furthermore, as is clear from fig. 4, the magnetic moment of a ferromagnetic at saturation must, on that theory, be equal to  $N\mu$ . In the polar theory, according to what has been said above, we have an essentially different situation.

Firstly, the paramagnetic (or diamagnetic) case happens much more frequently here; we can very well have paramagnetics (or diamagnetics) with a positive exchange integral. In particular, the metals with the number of free electrons equal to the number of atoms—the alkalis!—can never be ferromagnetic in our scheme, whatever the sign of the exchange integral. Ferromagnetism, in this scheme, results from the coincidence of a whole series of special conditions, corresponding, obviously, to what one would expect physically.

Secondly, we have here a natural explanation of the fractional values of the elementary magnetic moment of ferromagnetics. In fact, the case ( $\gamma$ ) of § 3 (*d*) corresponds to a ferromagnetic, whose magnetic moment at saturation is  $(N - 2s_0)\mu$  and not  $N\mu$  as with Heisenberg. For such a metal the elementary magnetic moment will seem to be

$$\mu' = (n - s_0) \mu/n,$$

*i.e.*, will have a fractional value. Experiment shows that this is realized in nature. The case ( $\alpha$ ) has never been observed experimentally.

### *Concluding Remarks.*

By the same method as for the energy, one can also calculate the values of the total current carried by the  $\Phi$  and  $\Psi$  waves of our scheme (the  $\phi$  and  $\psi$  waves, as mentioned in the introduction, give no current). The results of this calculation will be given in another paper, which will also contain the theory of the electrical conduction, *i.e.*, of the scattering of the  $\Phi$  and  $\Psi$  waves by the Debye waves of the lattice.

### *Summary.*

The most serious defect in the current theory of metals is that the models used do not allow of the simultaneous treatment of electric and magnetic properties, *e.g.*, conductivity and ferromagnetism. For example, in Heisen-



berg's model which can be applied with some success to ferromagnetism, the approximation on which the theory is constructed considers only *non-polar* states of the constituent atoms, and therefore no one of the states of the model can carry any current. The necessary generalization is attempted in this paper by including in the zero approximation of Heisenberg's model also *polar* states in which some of the atoms have two electrons and others none. The inclusion of these states allows for travelling electrons and therefore for states carrying current and also modifies in an important way the conditions for a ferromagnetic. These conditions prove to be much more restrictive than appears in Heisenberg's theory and ferromagnetics should be rarer than that theory leads one to expect—in conformity with the facts. The conditions for a ferromagnetic which replace the simple  $J > 0$  of Heisenberg's theory are formally set out here and appear to explain why the alkalis cannot be ferromagnetics even if they have a positive  $J$ .

---

*Statistical Measurements of Turbulence in the Flow of Air  
through a Pipe.*

By H. C. H. TOWNEND, D.Sc.

(Communicated by G. I. Taylor, F.R.S.—Received December 29, 1933.)

[PLATES 1-3.]

1. *Introduction.*

Numerous methods have been used to investigate experimentally the turbulence arising in a moving fluid from the boundaries confining it or the bodies immersed in it. Many of these only give qualitative information, whilst of those which yield numerical results most are rather incomplete in the amount of information they give about the distribution of the turbulent velocity components and their relation to one another.

In the present case the problem has been attacked by a method which enables small masses of air to be identified and their individual motions observed. Cinematograph records of these motions, and their subsequent analysis, allow quantitative results to be obtained in statistical form showing the magnitude of the mean turbulence at any point in a field of flow and its distribution over the field. The maximum values of the turbulent velocities and the average resultant velocity are also obtained in the process.

The method of identifying a small mass of air consists in producing a small electric spark between very fine electrodes situated at a point in the flow. When the spark occurs it heats up the air in its immediate neighbourhood, with the result that its refractive index is altered. By means of the Schlieren method of photography, this hot spot of air is rendered visible and its subsequent motions are photographed.

The possibility of using this method for measuring turbulence was suggested by some earlier experiments\* in which shadow photographs had been obtained of the filament of hot air in the wake of a fine platinum wire, heated electrically. If the flow is turbulent, a snapshot of such a filament shows breaks and changes of direction. Fig. 1, Plate 1, is a shadowgraph of the flow at the centre of a 4-foot wind tunnel at a speed of 22 ft./sec., obtained by placing a grid of heated wires across the stream; the wires were  $\frac{1}{2}$  inch apart, 0.002 inch diameter and 0.8 inch long.

Although a general idea of the turbulence can be obtained by means of hot wire filament lines they cannot be used to measure the turbulence quantitatively.† For the latter purpose it is necessary to photograph the movements of a large number of individual masses of air, preferably produced regularly at short intervals of time, from which the lateral and longitudinal velocity fluctuations at a fixed point may be obtained.

To explore the possibilities of such a method, a spark gap was set up on the axis of a 3-inch square tunnel which was fitted with glass sides. By means of suitable shutters, shadowgraphs of the motions of the hot spots produced by the sparks were obtained on photographic plates.

Fig. 2, Plate 1, is an example of records obtained in this way: on the left paths of particles are shown. (These must be distinguished from the "filament lines" obtained with a hot wire which they closely resemble in appearance.) On the right are snapshots of individual hot spots at successive points downstream.

The records marked (a) refer to the empty tunnel, while (b) and (c) show the turbulence 17 diameters behind a  $\frac{1}{4}$ -inch cylinder with its axis in the plane of the photograph. The wind speed was 34 ft./sec.

### *The Schlieren Method of Photography.*

The simple shadow method just described necessarily produces a magnified image of the field of flow and this involves the use of large plates if more than

\* 'Aero. Res. Ctee., Rep. Mem. No. 1349' (1930).

† They can, however, be used to indicate eddy frequencies. See p. 207.

a few inches of the field are to be covered. This difficulty entirely precludes the possibility of taking a rapid series of records, for which a cinema camera is essential. Since the hot spots are not self-visible, so to speak, they cannot be photographed directly, and hence to produce a record of small size it is necessary to collect the light after it has passed through the tunnel and render it convergent either by a lens or a concave mirror. The latter is most convenient on account of the large aperture required which must be large enough to include the whole field to be covered. It is also free from chromatic troubles.

Not only does this method enable the image of the field to be reduced to the size of a cinema film, but it also increases the intensity of the illumination enormously. It has the further advantage that it permits the use of the Schlieren method of photography which, being more sensitive than the shadow method, allows a smaller spark to be used and hence a smaller mass of hot air. Thus whilst fig. 2 was obtained with a spark gap of 5 mm. it was found possible with the Schlieren method to obtain satisfactory records with a gap of only  $\frac{1}{2}$  mm. Moreover, it was possible to photograph such a hot spot for a distance downstream several times greater than when using the shadow method.

Fig. 3, Plate 2, is a Schlieren photograph of the flow on the axis of a 3-inch pipe with a 1-mm. spark gap taken on a continuously moving film. The movement of the film causes the successive snapshots to be staggered along the length of the film. Here two views at right angles are shown.

#### *Description of Apparatus as used for Turbulence Experiments in a Pipe.*

In order to record both of the lateral velocity components simultaneously, which was essential if phase relations were to be obtained, it was necessary to divide the beam of light from the arc lamp into two pencils and to arrange that one of them should cross the pipe in a horizontal plane and the other in a vertical plane. After passing through the pipe both pencils were reflected from the spherical mirror into a cinema camera and photographed side by side simultaneously.

A sketch of the apparatus is shown in fig. 5. A 5-amp. arc lamp was focussed on the vertical edge of a diaphragm D. Close behind this diaphragm was a stroboscope disc S having 8 radial slits  $1^\circ$  in width and  $45^\circ$  apart. The emergent beam of light was divided into two flat pencils (a) and (b). These were passed through two identical shutters  $S_1$ ,  $S_2$ , consisting of flat vanes rotating on a common axis parallel to the wind, displaced axially so as to expose successive sections of the airstream in turn so that each flash from the disc S photographed the hot spot at a different point downstream. These shutters and the disc S

were coupled to the magneto and all were driven at the same speed. Thus in each half-revolution one hot spot was made, and was photographed four times while crossing the field. The pencil (a) passed horizontally through the air-stream and after reflection in the concave mirror M was brought to a focus on the diaphragm  $d_1$ .

The pencil (b) was made to cross the stream vertically by reflection in the two plane mirrors  $m_1, m_2$ , which were silvered on their front surfaces. This

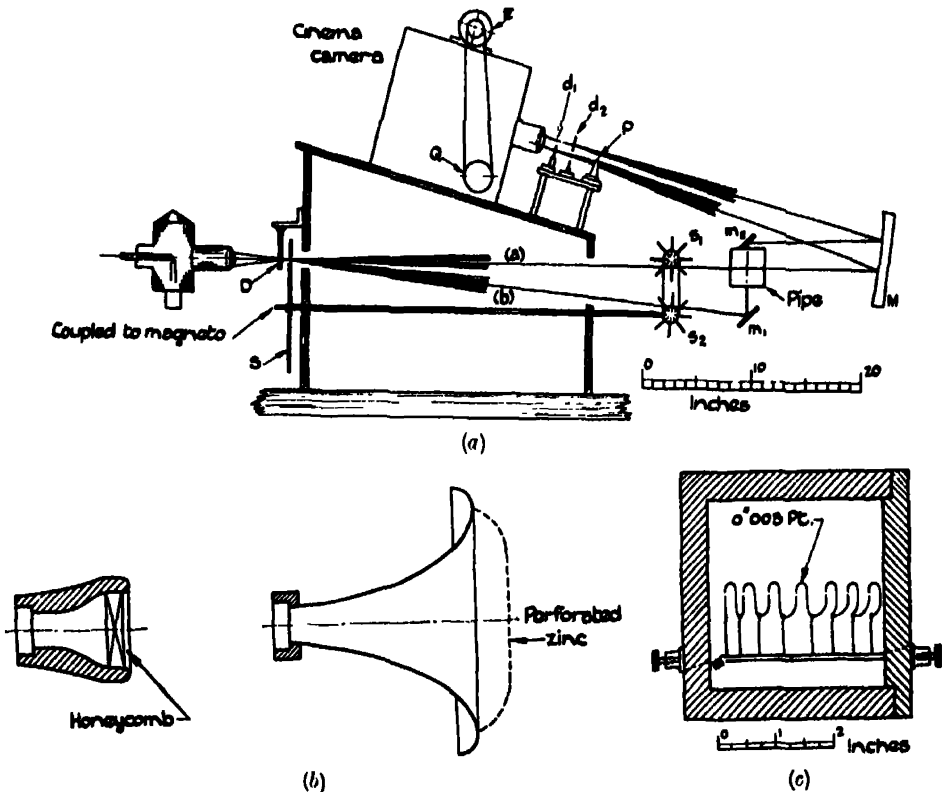


FIG. 5.—(a) Sketch of apparatus. (b) Small bellmouth; large bellmouth. (c) Arrangement of spark gaps in pipe.

pencil was then focussed on diaphragm  $d_2$ , which, owing to the longer path of the incident ray, was located slightly in front of  $d_1$ .

The slits  $d_1, d_2$ , were adjusted separately to intercept most of the light of their respective pencils.

Both pencils were then passed into a cinema camera and their images on the film were brought into proximity by interposing a prism, P, of suitable angle ( $7^\circ$ ) in the path of (a).

The mechanism for propelling the film through the camera intermittently was dismantled and the film was moved continuously by an electric motor E mounted on the camera. No blurring resulted from this method of use since the speed of the film was never greater and generally much less than that of the image of the hot spot across it. The exposure was about  $1/4000$  second.

In order to prevent overlapping of the pictures and to avoid wasting film, the camera had to be driven at a predetermined speed. To attain and maintain this speed rapidly, a stroboscope disc was mounted on the camera pulley Q and illuminated by an "Osglim" neon lamp supplied with alternating current (50 ~) from the lighting circuit. The speed of the magneto and its associated apparatus was controlled in the same way.

It will be seen from fig. 5 that to avoid interference between the incident and reflected beams from the concave mirror M an appreciable space must be left between it and the pipe. This seriously reduces the size of field that may be investigated. With an 8-inch mirror it was only possible to follow the motion of a hot spot for at most 5 inches downstream. When, at a later stage, the mirror was replaced by one of 12 inches diameter the optical system was modified to enable greater use to be made of the available aperture. A photograph of the apparatus thus modified is shown in fig. 6, Plate 3, and fig. 7, *a*, as arranged for one pencil of light. For two pencils, the optical system was as sketched in fig. 7, *b*.

In fig. 7, *a*, the arc is focussed on to the edge of a stainless steel mirror D which becomes effectively the first diaphragm of the Schlieren system. This is situated extremely close to the centre of curvature of the concave mirror. An inverted image of this edge M is received on the adjustable diaphragm *d* which is arranged to intercept most of the light as before. The remainder passes between the two edges, through the stroboscope disc S, and on into the camera.

Since the slit between the edges of the two diaphragms D and *d* may be reduced to very small dimensions (about 1 mm. in the present experiments) the light is reflected, almost perfectly, back along its incident path and hence passes through the hot air twice, which increases the visibility of the hot spot. Another advantage of this system is that the pipe may be brought very close to the mirror so that full advantage may be taken of its aperture. This in turn further improves the coincidence of the incident and reflected rays.

In this system it is necessary that both pencils of light should come to the same focus in order to use the same diaphragm.

*Application to the Measurement of Completely Turbulent Flow in Pipes.*

The present method of measuring turbulence has been applied to the flow of air through a pipe above the critical value of Reynolds' number, and as suggested by Mr. Fage, the conditions have been chosen to correspond with those prevailing in the experiments recently made at the N.P.L. with an ultramicroscope on the turbulent flow of water through a pipe.\* Since the results of the present work constitute an extension of those of the ultramicroscope experiments a brief résumé of the latter will be given.

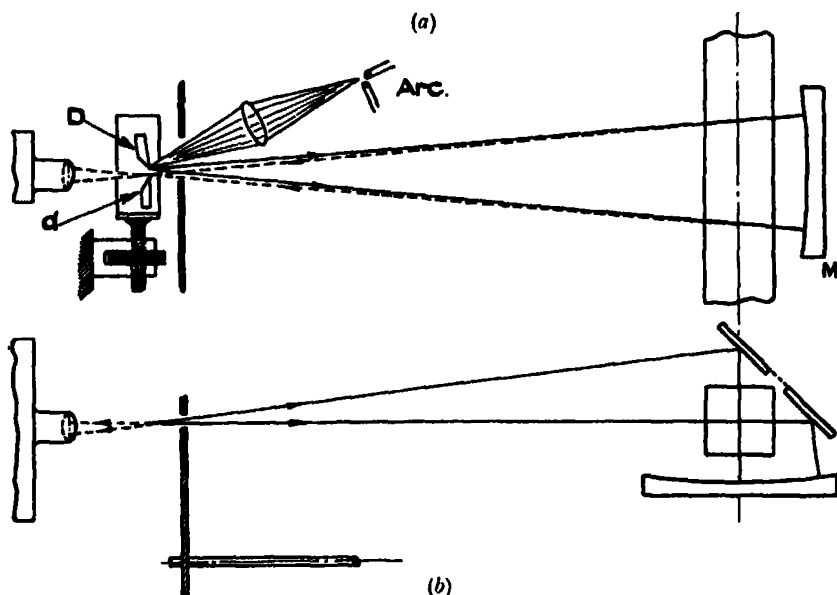


FIG. 7.—(a) Detail of source and diaphragms. (b) Arrangement when measuring  $u$ ,  $v$ , and  $w$ .

*Results of Ultramicroscope Experiments.*

In the experiments with the ultramicroscope the flow of water through a pipe of square section was studied by observing the motions of the minute particles of matter, which are normally present in suspension in tap water, under intense illumination. Owing to the extreme minuteness of these particles, the character of the turbulence very near the walls could be examined. In the present experiments the objects of observation are small elements of hot air whose boundaries are not clearly defined and no attempt has been made to use them to explore conditions very close to the surface. Except in this respect, however, the two experiments are closely related.

\* Fage and Townend, 'Proc. Roy. Soc.,' A, vol. 135, p. 656 (1932); and 'Aero. Res. Ctee. Rep. Mem., No. 1474' (1932).

In the ultramicroscope tests, measurements were made of the maximum inclinations of the paths of the particles to the axis of the pipe in two perpendicular planes parallel to the axis and at various positions across the section of the pipe. From the values found for these angles, together with measurements of the maximum and minimum axial velocities at any point obtained with a microscope having a moving objective, it was possible to obtain maximum values ( $u_1$ ,  $v_1$ , and  $w_1$ ) of the turbulent velocity components  $u$ ,  $v$ , and  $w$  in terms of the mean axial velocity,  $U$ , at any point in the pipe. These quantities were then proposed as a criterion of the degree of turbulence present at any point in the fluid over a long period of time, and by exploring at different points the distribution of turbulence over the section was determined. The results are shown in fig. 8.\*

### *The Present Experiments.*

Now in the experiments just summarized it was only found possible to measure the *maximum* values of the turbulent velocities. No information could be obtained about the mean turbulence. Further, although it was concluded that  $v$  and  $w$  were always out of phase with  $u$ , nothing more than this could be determined about any possible relationship that might exist between the phases.

The ability to produce a regular series of identified elements of the fluid which is a feature of the spark method provides the opportunity of measuring, not only the maximum, but all values of  $u$ ,  $v$ , and  $w$ , and in addition the phase relations between them, as will be described later.

It may be pertinent to mention here that one of the conclusions formed during the ultramicroscope experiments was that "in turbulent motion the fluid near the surface moves in relatively large masses." If this is so near the surface, it seems still more likely to occur in the main body of the fluid, and from this it would seem that the finite size of the hot spots produced by the sparks—which is presumably the chief weakness of the present method—may not be so serious as might at first appear, especially if the geometrical scale of the experiments is not too small.

### *The Measurement of the Mean Turbulence.*

In view of the foregoing, the main object of the present work has been to obtain a statistical estimate of the mean turbulence at any point in a pipe

\* Reproduced from Fage and Townend, 'Aero. Res. Ctee. Rep. Mem. No. 1474' (1932).

and its relation to the maximum value from a photographic record of a rapidly produced series of sparks, and to find if any correlation exists between the three components. The measurement of the maximum values also provides a means of checking the results against those of the ultramicroscope experiments.

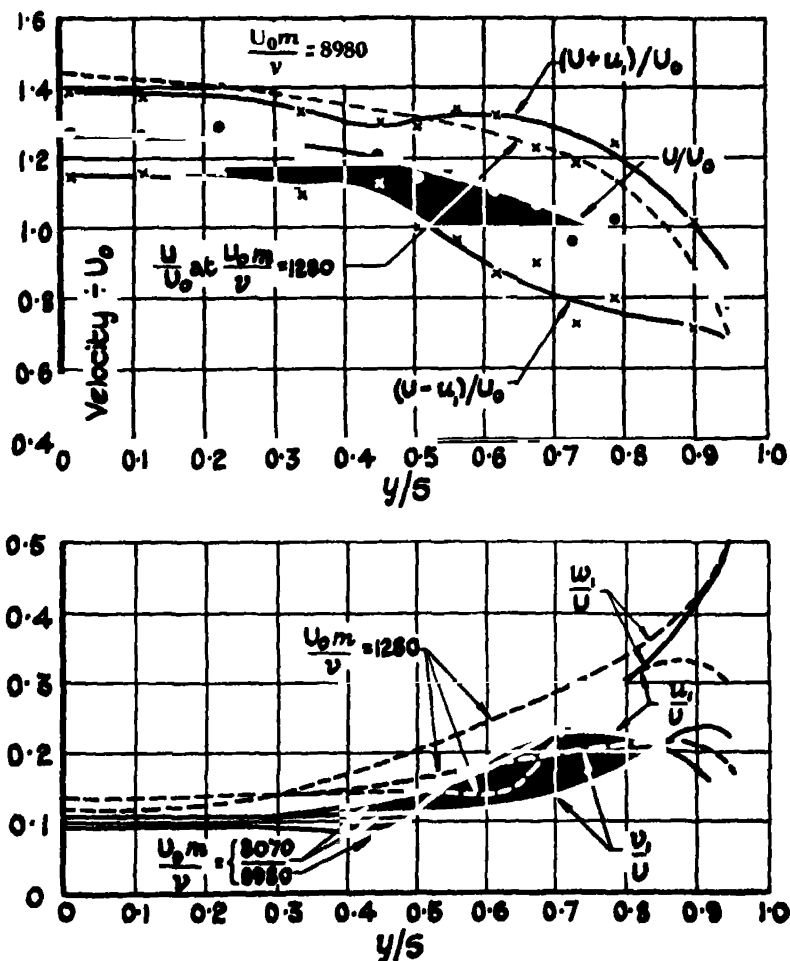


FIG. 8.

*Description of the Pipe.*

The pipe was of square section 3 inches by 3 inches. The working section was formed by four glass windows 11 inches long. The vertical walls were glazed for the full depth of 3 inches ; the windows in roof and floor were  $2\frac{1}{2}$  inches wide. Three lengths of pipe were provided, extending for 1, 28, and 60 diameters respectively from the entrance to the working section. Two kinds of



bell-mouth were used for the entry of the pipe. These are shown in fig. 5, b. The smaller bell-mouth was identical in form with that used in the ultramicroscope experiments except that in addition a honeycomb could be fitted if desired. The larger was used with the short pipe of 1 diameter. It was given a very high contraction ratio and fitted with a baffle of perforated zinc, shaped so as to be very roughly normal to the lines of flow of the entering air. This bell-mouth was used chiefly to provide an approximately non-turbulent stream for observing the character of the hot spot in the absence of turbulence. It may be noticed in passing that the larger bell-mouth was more effective than the smaller even with a honeycomb, in suppressing small scale turbulence, though the latter *when fitted with a honeycomb* of standard design gave a steadier distribution of velocity below the critical velocity.

The exit end of the pipe was enclosed in a box containing a centrifugal fan which drew air through the experimental pipe and discharged it to atmosphere. A speed of about 35 ft./sec. was obtainable.

#### *Measurement of Displacement of Hot Spots.*

To enable the displacement of the hot spots to be obtained scales were engraved on one vertical and one horizontal window which were photographed with the hot spots. They were calibrated by placing a scale along the axis of the pipe, and measuring the distance between the shadows of the graduations on the glass falling upon it.

#### *Choice of Reynolds' Number.*

In the ultramicroscope experiments it was found that for a square pipe of about 50 diameters length the flow is laminar below  $U_0 m/\nu = 500$ . The critical range extends to about 1400, but turbulence has not reached its full development until about 2000. In the present experiments this corresponds to a speed of about 7.2 ft./sec. at the centre (from fig. 9), but a higher speed than this was considered advisable partly to avoid appreciable convection due to buoyancy and partly because the hot spots are clearer at high speeds than at low. On the other hand, too high a speed necessitates very rapid cinematograph recording if large space intervals between successive hot spots are to be avoided. An axial speed of about 10 ft./sec. was adopted, giving a Reynolds' number of approximately 3000.

The results at this speed, however, suggested that the conditions at  $U_0 m/\nu = 3000$  were not comparable with those of the ultramicroscope. They were, therefore, repeated at the highest Reynolds' number in the experiments

made with the ultramicroscope although this meant using sparks more widely separated than at the lower Reynolds' number.

At the lower Reynolds' number the smallest interval between successive sparks on the axis was about 1 inch. Fig. 4, Plate 3, shows part of a film in which only one pencil of light was used, and the shutter  $S_1$ , fig. 5, *a*, was omitted. In fig. 4 the spark frequency is twice the picture frequency so that the first and second sparks in one picture have become the third and fourth in the next, the fifth and sixth in the next and so on. Each pair appears four times before going out of the field. The second and third sparks in the picture,

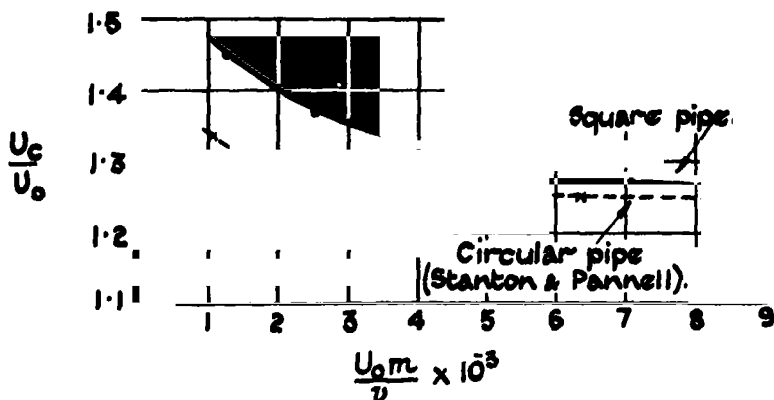


FIG. 9.—Variation of  $U$ ,  $u_1$ ,  $v_1$ ,  $w_1$  on the axis of the pipe with  $U_0 m / \nu$ .

marked A, if followed through the next three pictures, will be seen to have changed places so that with a field of view of this length closer spacing is not desirable.

### *The Wave-lengths of Particles.*

On the axis, the particle paths seldom show any curvature within the field of view, i.e., the wave-lengths are greater than four times the length of field. Near the wall they get rapidly smaller and it is necessary to measure velocities over much smaller distances to avoid curvature. Thus in measuring  $v$  at  $y/s = 0.8$ , if a straight line were drawn from the electrodes to a spark 1 inch downstream it would intersect the wall before reaching the next spark position 2 inches downstream, so that very close spacing was necessary in measuring the velocities. Wave-lengths have not been measured in these experiments since most of them are much longer than the field of view. Some idea of a possible lower limit to the wave-length, however, can be obtained from the signs of curvature which appear occasionally in the path of a particle. Thus, on the axis, out of several hundred paths photographed, only one or

two show signs of curvature developing before they leave the field of view 5 inches downstream from the electrode.

A further test of this point was made by inserting electrodes 5 inches upstream from the beginning of the window so that particle paths could be examined for curvature at distances of about 10 inches downstream. Here curvature was much more in evidence, and occasionally a spot on the axis which had moved some little distance laterally before entering the field had returned to the axis again just before leaving. This would correspond to a  $\frac{1}{2}$ -wave-length of about 10 inches or 3 diameters as a probable minimum on the axis, though it should be noted that in these experiments the view at right angles was not recorded.

At  $y/s = 0.8$  the  $\frac{1}{2}$ -wave-length of the component normal to the wall has a probable minimum of the order of one diameter.

#### *Measurement of the Mean Speed.*

For points off the axis of the pipe the spark record does not provide any means of ascertaining the axial speed or the mean rate of flow. It was, therefore, necessary to arrange a second spark gap on the axis which could be switched in for a short interval at the beginning and end of a run from which these speeds could be found. This was provided by arranging one of the supports carrying the platinum electrodes to be rotatable about its length and providing an extra pair of platinum electrodes at the axis of the pipe fixed to the same supports. These second electrodes were normally not opposite one another but by rotation could be made so; this, of course, simultaneously separated the other electrodes constituting the gap situated at the point under examination so that they did not spark. The spark was therefore easily transferred from one position to the other, and as the whole experiment only lasted some 20 seconds, the air stream could be relied on to maintain a steady velocity.

#### *The Instantaneous Velocity Distribution.*

In order to study how the velocity distribution across the pipe varies with time, a row of seven spark gaps was arranged across the pipe as shown in fig. 5, c. A streamlined fibre support  $1/20$  inch thick was fixed across the pipe near one wall, and 0.01 inch steel wires projected towards the centre carrying platinum electrodes 0.003 inch thick. These were arranged so that the spark gaps were about  $\frac{1}{2}$  mm. long, in series, one being on the axis and the others symmetrically on each side. The two extreme gaps were  $\frac{1}{4}$  inch from the

walls. In these experiments it was necessary to boost the voltage of the magneto since the seven gaps required a considerably greater voltage than it normally produced. This was effected by applying extra magnets to the magneto and inserting a 10-volt battery in series with the primary. By this means sparks of one polarity had their voltage augmented at the expense of those of opposite polarity, which were suppressed altogether. The safety gap was also increased.

*Axes of Reference.*

The axis of the pipe is taken as  $x$ -axis, the  $y$ -axis is horizontal and perpendicular to one side of the pipe, and the  $z$ -axis is vertical.

*List of Symbols.*

$2s$  = length of side of pipe = 3 inches.

$m$  = hydraulic mean depth =  $s/2$ .

$U$  = mean velocity at any point in the pipe.

$U_0$  = mean rate of flow through pipe.

$U_a$  = mean velocity on the axis.

$u, v, w$  = components of turbulent velocity.

$u_1, v_1, w_1$  = maximum values of  $u, v$ , and  $w$ .

$\bar{u}, \bar{v}, \bar{w}$  = arithmetical average values, neglecting signs, of  $u, v$ , and  $w$ .

$t$  = time taken for hot spot to reach  $(x, y, z)$ .

$N$  = total number of observations in one record.

$f$  = frequency of occurrence of a given value of  $u, v$ , or  $w$ .

$i$  = group interval used in the histograms of  $\frac{U, \&c}{U_0} = 0.02$  throughout the report.

*Method of Analysing Films.*

At each point the velocities were derived from a film having usually about 200 pictures. The film was projected on to a sheet of squared paper at a suitable magnification and for each spot values of  $x, y$ , and  $z$  were read off at a position corresponding to a fixed interval of time,  $t$ , from its origin which was adjusted to occur very close to the electrodes.

The velocities at the point were therefore

$$U = \frac{1}{t} \cdot \frac{\sum x}{N} = \frac{\bar{x}}{t} = \text{mean speed,}$$

$$u =$$

$$v = \frac{y}{t} \quad : \text{turbulent components,}$$

$$w = \frac{z}{t}$$

while the velocity on the axis was  $U_0 = \frac{\bar{x}_0}{t}$ , so that

$$\frac{u}{U_0} = \frac{x - \bar{x}}{\bar{x}_0},$$

$$\frac{v}{U_0} = \frac{y}{\bar{x}_0},$$

$$\frac{w}{U_0} = \frac{z}{\bar{x}_0}.$$

The time interval  $t$  only entered directly into the calculation of mean speed and Reynolds' number. The value of  $U_0$  was inferred from the Reynolds' number and the curve of fig. 9.

The velocity on the axis  $U_0$  has been used throughout instead of  $U_0$ , the mean rate of flow, because it is the quantity which has been measured. To measure  $U_0$  in a square pipe with air as the working fluid would require a velocity exploration across the whole section; the distribution also changes with Reynolds' number, as shown by the curve of  $U_0/U_0$  given in fig. 9. The value of  $U_0$  has, when required, been taken from this curve, and so has the Reynolds' number.

During the analysis a watch was kept for signs of curvature in the paths of the spots.

The results were then analysed statistically to ascertain whether the velocities were distributed at random. The root-mean-square value,  $\sigma$ , of each of the components  $u/U_0$ ,  $v/U_0$ , and  $w/U_0$  was found from the observations, and a frequency curve calculated from the formula

$$\frac{f}{N} = \frac{i}{\sigma \sqrt{2\pi}} e^{-\frac{(\frac{u \& v}{U_0})^2}{2\sigma^2}},$$

where  $f/N$  is the fraction of the total number of observations having a specified

value of  $u/U_0$ . The observations were then arranged in suitable groups and a histogram plotted on the same diagram as the frequency curve. The degree of conformity between the two curves shows the extent to which the normal distribution curve represents the observations. In discussing the results  $u_\sigma/U_0$  will be used to denote the value of  $\sigma$  obtained from the observations of  $u/U_0$  and similarly for  $v_\sigma/U_0$  and  $w_\sigma/U_0$ .

*Number of Observations (N).*

The number of hot spots observed in any one test was kept as low as possible on account of the labour involved in analysing a large number of films. As a rule about 200 observations were made of each quantity measured. One film having about 400 pictures was analysed in two parts and the values of  $\sigma$  from the first and second parts compared with each other and with the values obtained from the 400 taken together. Table I shows the result.

Table I.

	$u_\sigma/U_0$	$v_\sigma/U_0$	$w_\sigma/U_0$
First 200 observations	0.072	0.049	0.052
Second 200 observations	0.072	0.047	0.056
All 400 observations	0.072	0.048	0.055

The agreement of the results of each section of the film with that of the whole is fair and was thought to show that 200 observations would give satisfactory accuracy.

*Accuracy of the Measurements.*

(a) *Average Values of  $u$ ,  $v$ , and  $w$ .*—The accuracy with which the position of a hot spot can be read from the film is rather low for two reasons. The image of the spot is rather ill-defined except very close to the electrodes. Nevertheless in streamline motion the centre of the spot can be located with reasonable accuracy (about 0.02 inch) several inches downstream. In turbulent motion, however, the spot is often rapidly distorted and sometimes broken up into two or even more discrete fragments. Usually the "centre" of the distorted mass has been estimated as far as possible, but when two distinct parts have been discernible, in *both* views, they have been recorded as double readings and counted as two observations.

In finding mean values, where large numbers of spots are measured, the final value is unlikely to be much affected by errors due to the above causes, since they will probably cancel out.

(b) *Maximum Values,  $u_1$ ,  $v_1$ , and  $w_1$ .*—With maximum values the accuracy is very much lower than for mean values since the result depends on one or two solitary observations. The error in estimating the position of a single spot may be as much as 5 or even 10%, although most of the maxima recorded are certainly more accurate than this.

Again, although most of the films contained about 400 observations from which to select maxima, this number would seem inadequate for the purpose of finding the maxima with any certainty since it may lead to underestimation of the maxima, but cannot lead to overestimation.

In the ultramicroscope experiments, observations of the angular deviations to find  $v_1$  and  $w_1$  was maintained for 15 or 20 minutes, during which time multitudes of particles must have been observed. The observations with the rotating objective (to find  $u_1$ ) were intermittent as the field of view was only visible for a small part of a revolution, so that the number of particles observed in a given time was considerably smaller. In the present tests the record seldom corresponded to more than 10 seconds, though probably a much higher proportion of the observations were effective. The values of  $u_1$ ,  $v_1$ , and  $w_1$  given in this report, while good as observations, may therefore sometimes give maxima which are somewhat too low.

In the ultramicroscope tests, the maxima were obtained by taking the mean of the positive and negative readings since there was no method of measuring the mean, which was in general different from zero on account of the secondary motions present in a square pipe. In the present work, however, the mean was observed, and the maxima recorded are the greatest deviations from the mean irrespective of sign. They usually exceed the mean values of the two greatest deviations of opposite sign sometimes by 10%, as would be expected with only a small number of observations.

For the component parallel to the wall, the average value was always positive, on account of the buoyancy of the hot spot. This component of the secondary motion should be zero in a horizontal plane. For the component normal to the wall, the average value was always negative, i.e., it revealed the secondary motion directed from the centre of the wall towards the axis, but was unaffected by buoyancy which acted at right angles to the plane of observation.

### *Results.*

The first test carried out was made with the object of comparing the values of  $u_1$ ,  $v_1$ , and  $w_1$  with those found with the ultramicroscope. This was done

on the axis of the pipe, and, as mentioned above and shown in Table II, the agreement was good (though the number of observations was less than 200).

Table II.

	Ultramicroscope ( $U_0 m/\nu = 1280$ ).	Sparks ( $U_0 m/\nu = 2800$ ).
$\theta_{xy}$ $\theta_{xz}$ $u_1/U_0$	$6.7^\circ$ $0.114$	$6.3^\circ$ $5.7^\circ$ $0.115$

Here the conditions were similar to those in the ultramicroscope tests, including the shape of the bell-mouth, except that the working section was only 28 diameters from the entry of the pipe in the present case instead of 50 diameters.

It is known that after about 20 diameters from the entrance of a round pipe the velocity distribution across the pipe undergoes only slight change. However, to test whether the entry conditions produced any perceptible influence on the turbulence at 28 diameters downstream, a further record was taken with the large bell-mouth shown in fig. 5 *b*. The two tests are compared in figs. 10 (*a*) and (*b*). In the latter case (*b*) there is an appreciable reduction in the values of  $\sigma$  showing that the stabilizing influence of the large contraction is still felt at this distance downstream.

It should be noticed that this refers only to the flow on the *axis*. Later,\* when the same conditions were observed visually with seven spark gaps simultaneously it was noticed that the flow on the axis was appreciably steadier than elsewhere.

The pipe was then lengthened to 60 diameters, and two tests were made (on two components only,  $u$  and  $w$ ), one with the original small bell-mouth (no honeycomb) and one without the bell-mouth, i.e., with a bluff entry. For the latter test the frequency distributions are shown in fig. 10 (*c*) while the actual values of  $u/U_0$  and  $w/U_0$  are plotted in fig. 11. The values of  $\sigma$  for these two conditions agreed almost exactly, so that in turbulent flow the influence of the entry appears to be inappreciable at 60 diameters downstream. All the subsequent work was therefore done with this length of pipe and with a bluff entry. The values of  $\sigma$  at 60 diameters are somewhat smaller than at 28, with small bell-mouth at entry in each case.

\* See p. 207 and fig. 17 (*d*), Plate 3.



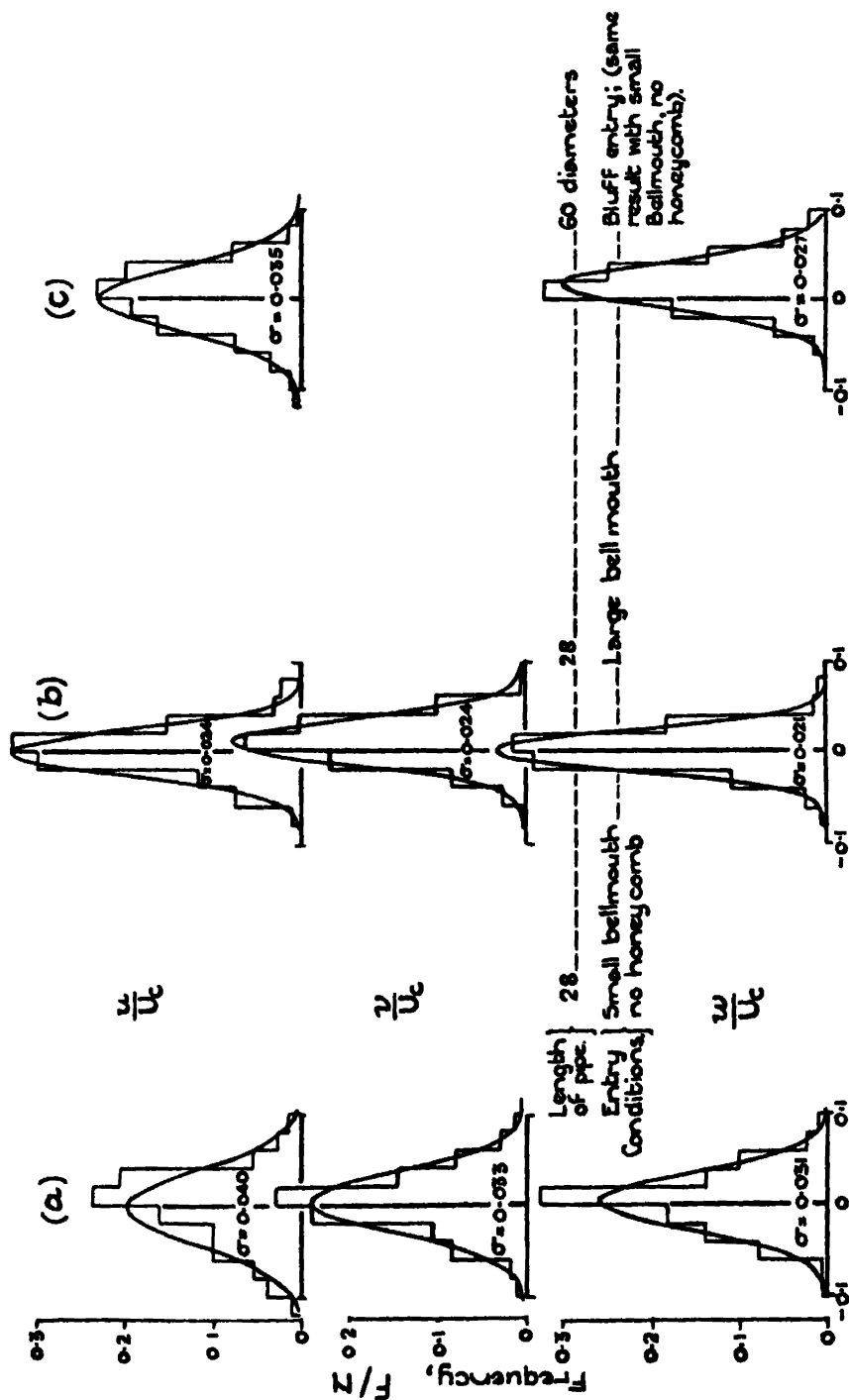


FIG. 10.—Turbulence on the axis of a square pipe. Influence of length and of entry conditions.  $U_{pm}/v = 2800/3500$ ,  $\sigma = \text{root mean square value of } u/U_p$ , etc.

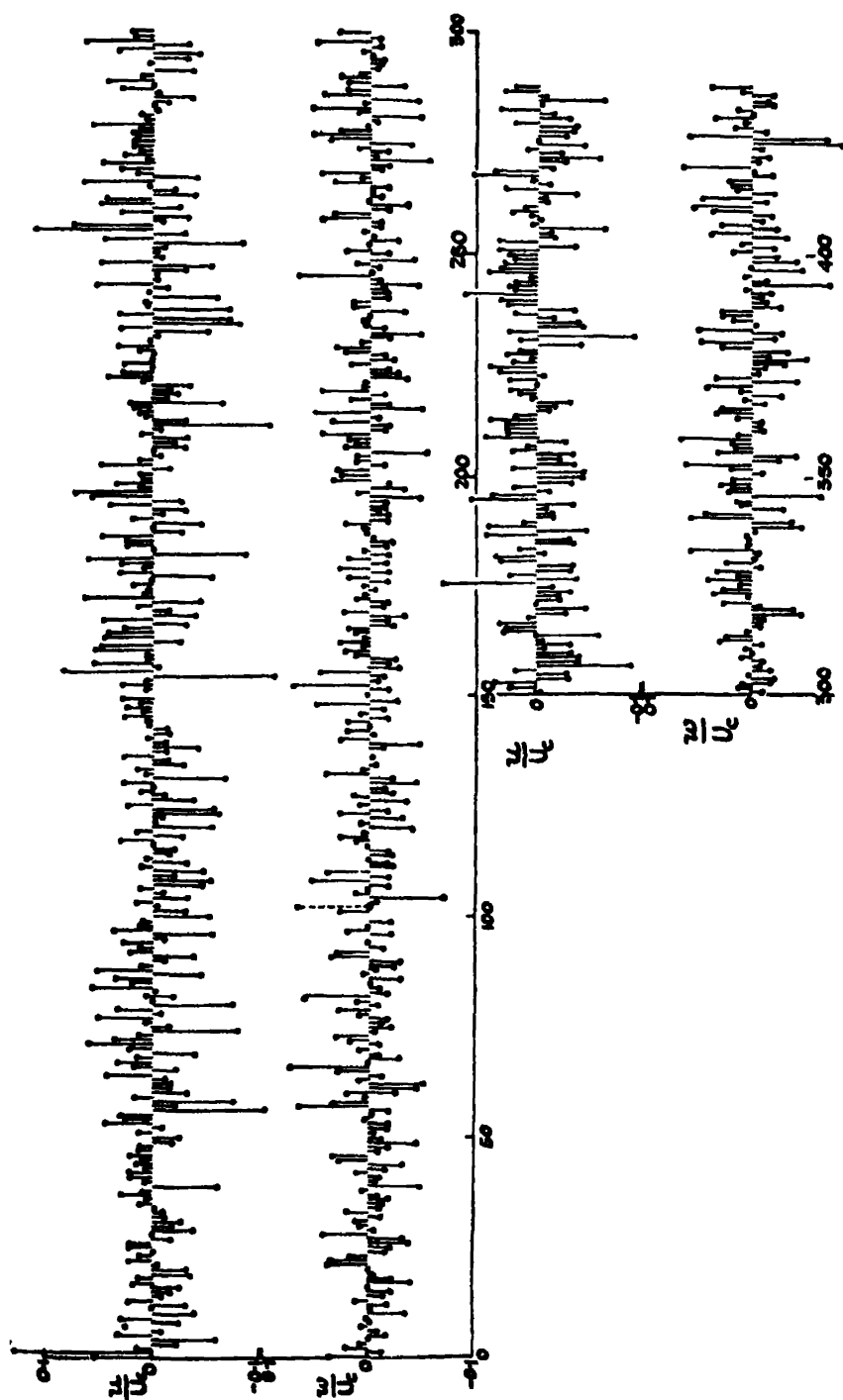


FIG. 11.—Turbulence on the axis of a pipe (with sharp entry) 60 diameters downstream. Simultaneous values of  $u/U_c$  and  $w/U_c$ .  $U_c = 12.3$  feet per second.  $U_{c,m}/y = 3600$ .

The change in turbulence for the two entry conditions at 28 diameters is shown also in figs. 12 and 13, where the velocities  $v/U_0$  and  $w/U_0$  have been plotted against  $u/U_0$  in correlation diagrams. The number of observations is different in the two conditions (183 and 460) which somewhat obscures the comparison, but the scattering is seen to be much less when the large bell-mouth is placed at the entry of the pipe. None of the diagrams shows any sign of a relation between the turbulent components.

### *Distribution of Turbulence across the Pipe.*

The results for both Reynolds' numbers are collected in Table III. Maximum, average, and root-mean-square values of the turbulent components  $u$ ,  $v$ , and  $w$  are given for various positions across the pipe.

For the higher Reynolds' number ( $U_0 m/\nu = 9300$ ) the frequency distributions for the turbulent components are plotted as histograms in fig. 14. For each histogram the value of  $\sigma$ , the root mean square of all the observations, measured from the mean value, is marked on the diagram. The smooth curves for  $u/U_0$  are calculated, using these values of  $\sigma$ , from the formula

$$\frac{f}{N} = \frac{i}{\sigma \sqrt{2\pi}} e^{-\frac{(u/U_0)^2}{2\sigma^2}},$$

and similarly for  $v/U_0$  and  $w/U_0$ .

The goodness of the fit of the two sets of diagrams indicates to what extent the observations are represented by the "error" law.

The frequency curves for the lower Reynolds' number (3000) are not reproduced, as they show similar characteristics. The use of  $\sigma$  to symbolize turbulence is particularly appropriate as its square is proportional to the energy at a point in the fluid due to one component of the turbulent velocity.

The total turbulent kinetic energy is proportional to

$$(q_\sigma/U_0)^2 = (u_\sigma/U_0)^2 + (v_\sigma/U_0)^2 + (w_\sigma/U_0)^2,$$

where  $q^2 = u^2 + v^2 + w^2$ .

The fraction of the total kinetic energy which is represented by turbulence is, at any point

$$\frac{(q_\sigma/U_0)^2}{(U/U_0)^2 + (q_\sigma/U_0)^2}.$$

The variation across the pipe in the turbulent kinetic energy at  $U_0 m/\nu = 9300$  is given in Table IV.

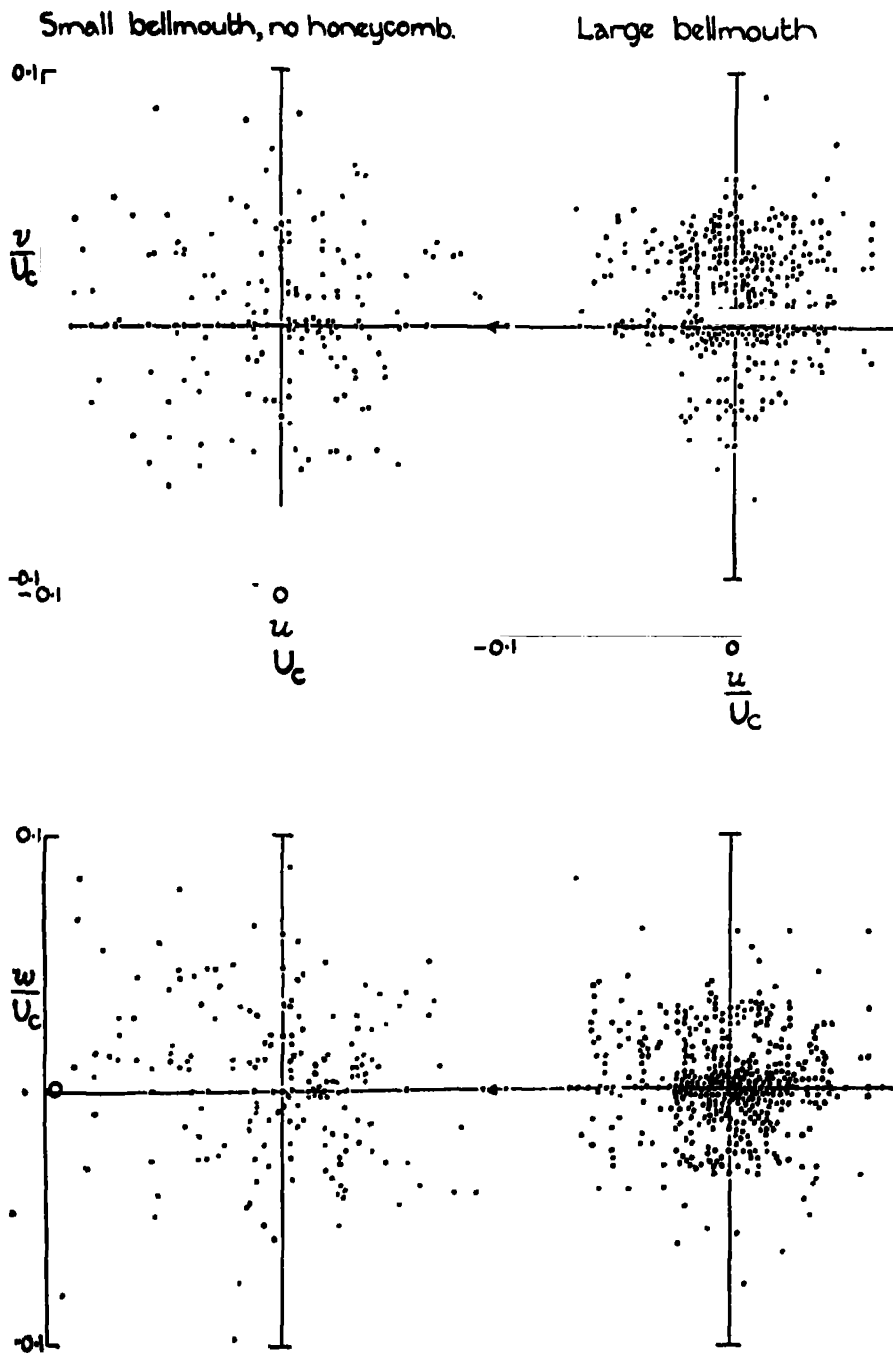


FIG. 12.

FIG. 13.

FIGS. 12 AND 13.—Turbulence on the axis of a pipe 28 diameters downstream. Correlation diagrams for  $u$ ,  $v$ , and  $w$   $U_0 m/\nu = 2700$ .

Table III.—Distribution of Turbulence in a Square Pipe.

$u_e/U_e$ , etc., is the value of the "Standard Deviation"  $\sigma$  for  $u/U_e$ , etc., in the frequency formula  $\frac{f}{N} = \frac{i}{\sigma\sqrt{2\pi}} e^{-\frac{(u/U_e)^2}{2\sigma^2}}$ .

It is the root-mean-square value of the observations.

Reynolds' number $U_0 m/\nu$	$y/s$	$U_e$	Maximum values.			Average values.			Root-mean-square values.			(Max.) $\tau$ (R.M.S.).			$u_1/u_e$	$v_1/v_e$	$w_1/w_e$	
			$u_1/U_e$	$\tau_1/U_e$	$w_1/U_e$	$\bar{u}/U_e$	$\bar{v}/U_e$	$\bar{w}/U_e$	$u_e/U_e$	$v_e/U_e$	$w_e/U_e$	$u_1/u_e$	$v_1/v_e$	$w_1/w_e$				
3000	0	9.61	0.121	0.106	0.122	0.032	0.028	0.026	0.040	0.034	0.031	3.0	3.1	3.9	Maxima from ultra-microscope at $U_0 m/\nu = 8000/9000$ + root mean square from spark tests at $U_0 m/\nu = 9300$ .			
	0.4	8.81	0.161	0.126	0.161	0.046	0.026	0.032	0.056	0.034	0.039	2.9	3.7	4.1				
	0.6	9.55	0.174	0.156	0.181	0.059	0.038	0.044	0.072	0.048	0.055	2.9	3.3	3.3				
	0.8	10.07	0.254	0.195	0.202	0.082	0.054	0.049	0.098	0.065	0.061	2.6	3.0	3.3				
9300	0	29.9	0.104	0.108	0.081?	0.028	0.023	0.024	0.035	0.030	0.028	3.0	3.6	2.9?	2.7	3.6	3.9	
	0.4	30.1	0.151	0.106	0.092?	0.044	0.027	0.028	0.055	0.032	0.034	2.7	3.3	2.7?	1.4	3.4	3.6	
	0.6	30.0	0.217	0.103	0.151	0.060	0.032	0.039	0.074	0.040	0.048	2.9	2.6	3.1	2.3	2.7	3.2	
	0.8	29.9	0.268	0.144	0.234	0.087	0.040	0.052	0.107	0.049	0.067	3.0	2.9	3.5	1.6	2.8	3.5	

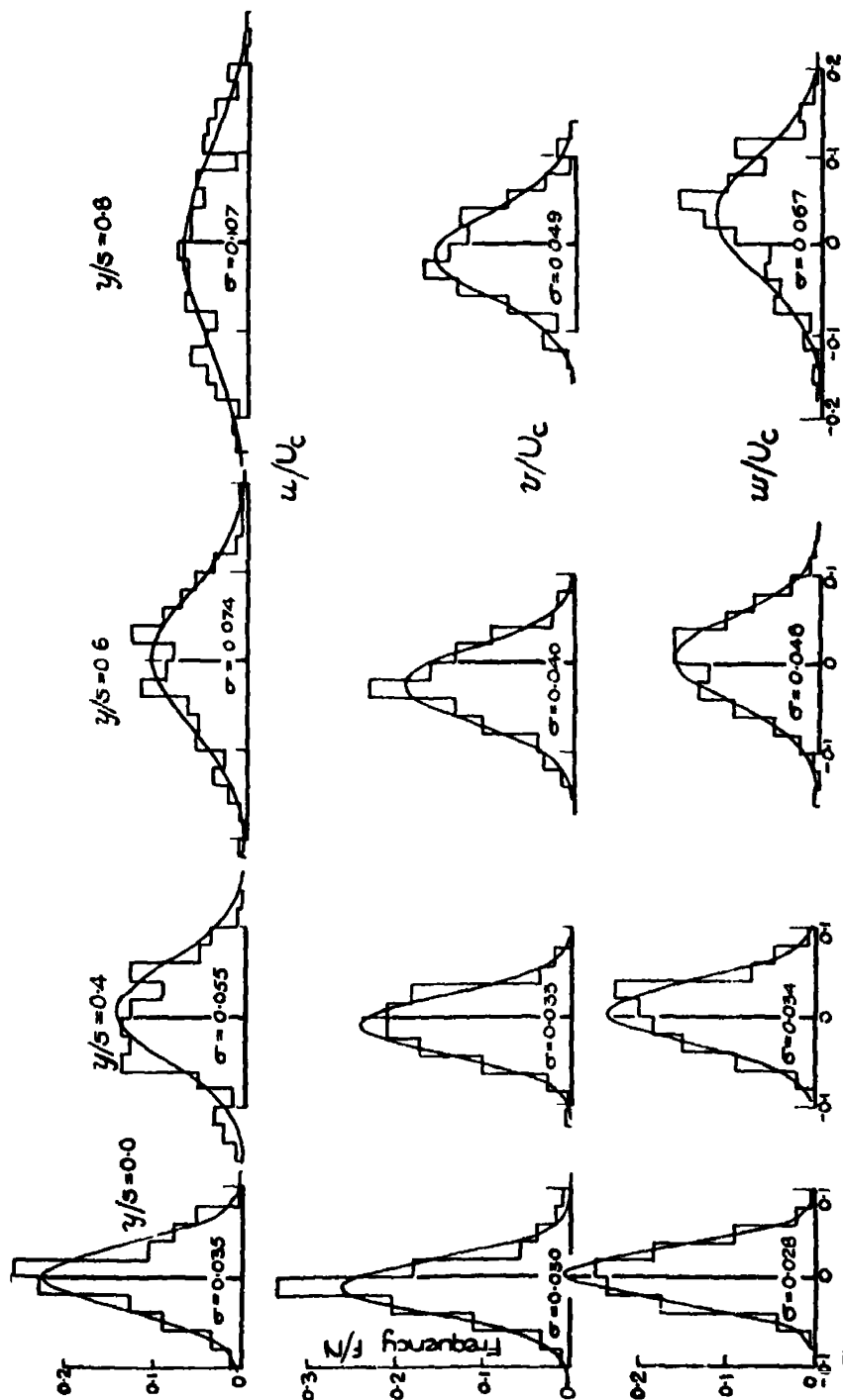


Fig. 14.—Distribution of mean turbulence across a square pipe.  $U_m/\nu = 3300$ .  $\sigma$  = root-mean-square value of  $u/U_c$ , etc.

Table IV.

$y/s$ .. . . .	0.0	0.4	0.6	0.8
$(q_\sigma/U_0)^2$	0.0029	0.0052	0.0094	0.0183
$\frac{(q_\sigma/U_0)^2}{(U/U_0)^2 + (q_\sigma/U_0)^2}$	0.0029	0.0057	0.0120	0.0289

Thus at  $y/s = 0.8$  the turbulent energy is only 3% of the total, or 2% of the total energy on the axis. The variation of  $(q_\sigma/U_0)^2$  across the pipe is shown by the dotted line in fig. 15.

In fig. 15 are also plotted the values of  $\sigma$  and of the maximum for each component for both Reynolds' numbers. Smoothed curves of the maxima from the ultramicroscope experiments are also shown corresponding to the higher Reynolds' number. As already mentioned the maxima obtained in the present experiments are liable to underestimate the maximum turbulent components, on account of the small number of hot spots in the records. This probably accounts for the low value of  $w_1$  near the axis which should by symmetry equal  $v_1$ , specially as at  $U_0 m/\nu = 3000$   $w_1$  is rather high. However, the fact previously observed,\* that the central region of equality of  $u_1/U_0$ ,  $v_1/U_0$ , and  $w_1/U_0$  extends to a greater distance from the axis at the higher Reynolds' number, is confirmed. For the values of  $\sigma$ , this effect is scarcely perceptible.

#### *Comparison with the Ultramicroscope Results.*

The agreement of  $v_1$  and  $w_1$  with the ultramicroscope results is good except for the low values of  $w_1$  near the axis which in the present case are probably underestimates. For  $u_1$ , however, whilst the agreement is good on the axis, the present results show much greater values at other points in the pipe. It is possible that higher values of  $u_1$  might have been obtained with the ultramicroscope if observation had been continued for a longer time to compensate for the reduction in effective time owing to the small fraction of a revolution of the rotating objective during which particles were under observation.

#### *The Validity of the Use of the Maxima as a Criterion of Turbulence.*

One of the objects of the present work was to discover whether the maximum values  $u_1$ ,  $v_1$ , and  $w_1$  were proportional to the mean turbulence at all points

\* Fage and Townend, 'Aero. Res. Ctee. Rep. Mem., No. 1474,' (1932).

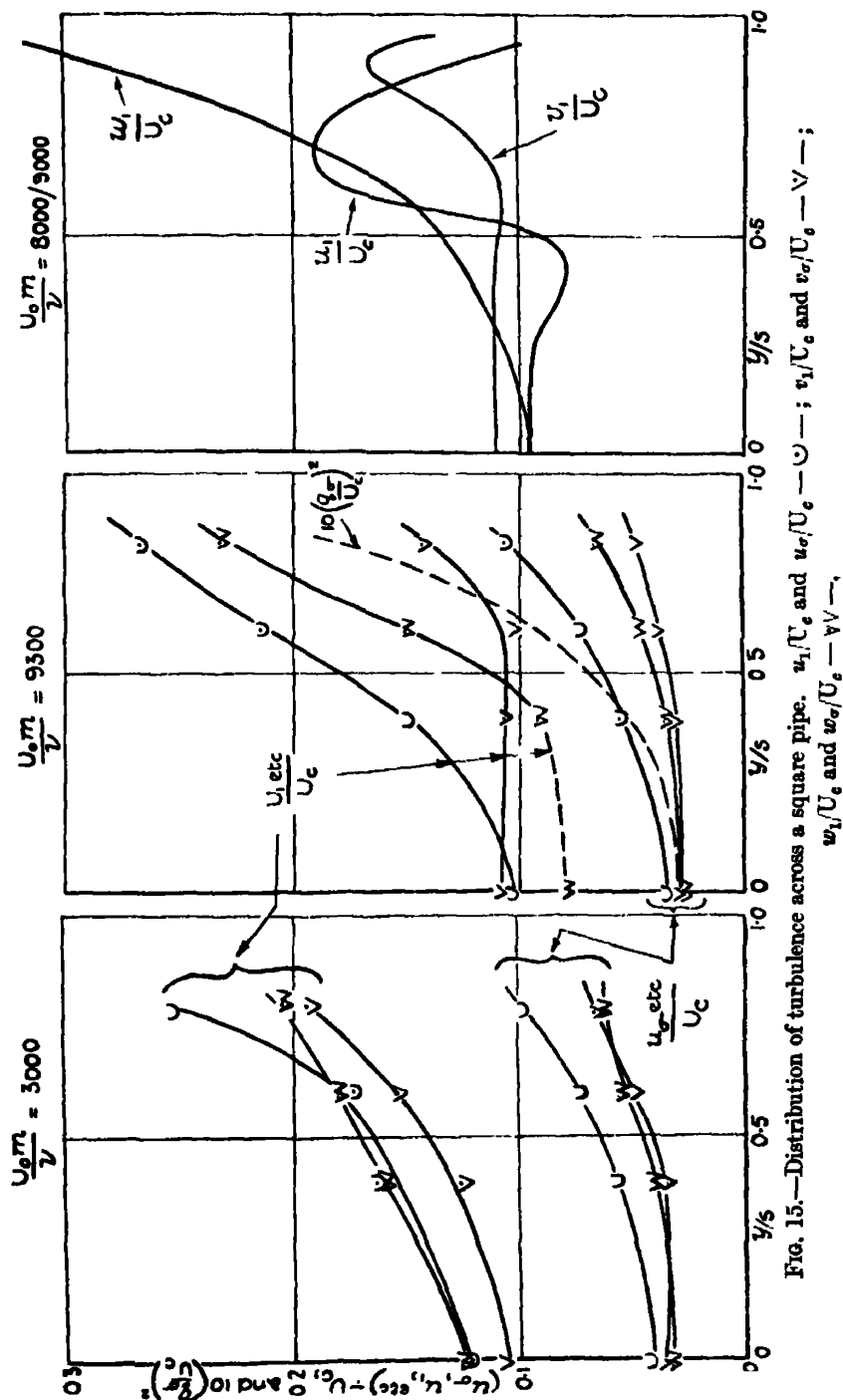


FIG. 15.—Distribution of turbulence across a square pipe.  $u_1/U_0$  and  $u_2/U_0$  — — —;  $v_1/U_0$  and  $v_2/U_0$  — — —;  $w_1/U_0$  and  $w_2/U_0$  — — —.



across the pipe, and could therefore be regarded as truly representative of the turbulence. The ratios of the maximum to the root-mean-square values are given in Table III.

They are somewhat irregular on account of the uncertainty of the maxima, but show that the maxima are roughly three times the root-mean-square values for all three components. The variation with distance from the axis is small, but suggests a slight tendency to fall as the wall is approached.

Table III also gives the absolute average values  $\bar{u}/U_0$ , etc., of the components. These are plotted in fig. 16 (a) against the corresponding values of  $\sigma$  and give the straight line  $\bar{u}/u_\sigma = 0.812$ . The theoretical value of the ratio average/r.m.s. for the "error" curve is  $\sqrt{2/\pi} = 0.798$ . A comparison of these two values is then an alternative method of estimating the accuracy with which the error law represents the observations. It has the advantage of showing all the observations on one diagram.

From the above results it is concluded that the values of the maxima represent the turbulence at a point in a pipe reasonably well, and that their values are roughly three times those of the root-mean-square or 3.75 times those of the average values.

It may be noticed that  $u_\sigma$  is everywhere consistently greater than  $v_\sigma$  and  $w_\sigma$ , presumably owing to the absence of any walls at right angles to  $u$  which can exert upon it a constraint analogous to that which the walls exert on  $v$  and  $w$  even at the centre of the pipe. This suggests that if the section of the pipe instead of being square were of rectangular section, very long in the direction of  $y$  compared with  $z$ , so as to approximate to two-dimensional mean motion the component  $v_\sigma$  would tend to equality with  $u_\sigma$  rather than with  $w_\sigma$ .

#### *Correlation between the Turbulent Components.*

The correlation diagrams in figs. 12 and 13 suggest that the components do not bear any relation to one another. A few values of the correlation coefficient

$$r = \frac{1}{u_\sigma v_\sigma} \cdot \frac{\Sigma (u \cdot v)}{N}$$

have been worked out for the components  $u$  and  $v$  at different points across the pipe. In general the correlation coefficient may vary from zero, indicating no correlation, to unity which indicates perfect correlation.

The results at  $U_0 m/v = 9300$  are given in Table V.

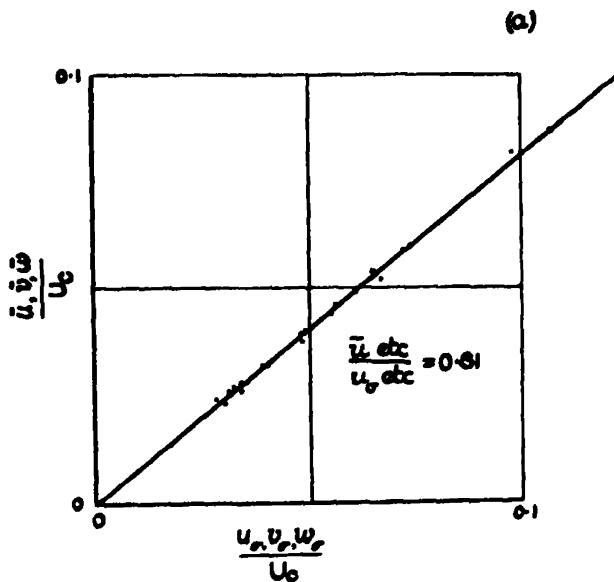


FIG. 16.—(a)

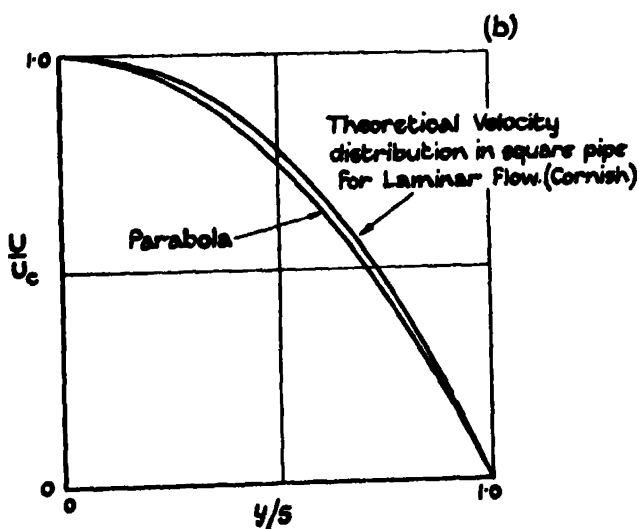


FIG. 16.—(b) Theoretical velocity distribution in square pipe for laminar flow (Cornish).

Table V.—Correlation between  $u$  and  $v$ .

$y/s$	0.0	0.4	0.6	0.8
$r$	0.05	0.17	0.16	0.18
$P$	$0.4 < P < 0.5$	—	$0.02 < P < 0.03$	$< 0.01$

$P$  is the probability that the observed values of  $r$  should occur in a large number of pairs of uncorrelated quantities.\* Hence the correlation observed at points off the axis is probably significant.

\* The values of  $P$  are taken from a Table of Significance (Table IV, p.137) of Fisher's "Statistical Methods for Research Workers" (1925).

### *The Instantaneous Velocity Distribution.*

Cinematograph records have been taken of the velocity distribution in a vertical plane across the centre of the pipe for various conditions of flow both above and below the critical Reynolds' number. Some examples of these records are reproduced in fig. 17, Plate 3. The fibre support carrying the electrodes can be seen on the right-hand side of each record. There are seven spark gaps and in each record the sparks are photographed almost at the instant of their occurrence. The mean axial speed  $U_0$  and the number of sparks per second is shown beneath each record.

The first two records are taken under the same conditions, and correspond to the lower Reynolds' number (about 3000) for which observations were analysed. They differ only in the fact that (a) was taken with the Schlieren slit horizontal while (b) (and also all the others) were taken with the Schlieren slit vertical. Since "striæ" are best revealed when they are parallel to the slit, the effect of having the slit horizontal is that the continuous filaments of hot air streaming from the heated electrodes themselves which act as hot wires are shown in addition to the hot spots due to the sparks themselves. This is at the expense of definition of the hot spots themselves, and for this reason the vertical slit has been used for all the other experiments.

The velocity distribution is best judged from the appearance of the first column of spots which are clear of the electrodes. Any spot in one picture has moved two places to the left in the picture below it. No quantitative measurements have been made from these films, but they serve to illustrate several features of interest.

Figs. 17 (a) and (b) show the rapidly fluctuating character of the velocity distribution, which may vary from an almost uniform distribution, as shown, for example, in the first picture in (a) to one of very irregular shape such as in

the following picture. From the filament lines in such a record it appears that eddies of fairly regular periodicity sometimes show themselves. Many of these have wave-lengths well under  $\frac{1}{2}$  inch corresponding to a frequency of about 300 per second. In the first picture of (b) the spot just below the axis shows a complete cycle with a frequency of 500 per second.

Fig. 17 (c) showing flow below the critical will be referred to later.

In fig. 17 (d) and (e) the pipe was 28 diameters long, and the influence of the large bell-mouth is still appreciable. In (d) the flow is turbulent at all points except on the axis. If the axial spot is examined it will be seen to repeat its position in each picture very closely while the other spots show turbulence which increases towards the wall. Visual observation of this condition showed that this characteristic was maintained for a considerable time. It corresponds to the case analysed in fig. 10 (b) and shows that the reduction in  $\sigma$  compared with fig. 10 (a), which was obtained with the small bell-mouth, probably only applies to the axial region of the pipe as mentioned earlier.

Fig. 17 (e) shows the effect of reducing the speed in (d) from 11.6 to 5.6 ft./sec. Here the distribution is nearly uniform and the flow practically non-turbulent.

#### *The Flow near the Critical Velocity.*

At speeds down to about 5 ft./sec. the effect of the buoyancy of the hot spots was almost negligible and it was thought that if allowance was made for this effect, which can easily be measured, the flow over the critical range and somewhat below it could be studied, and in particular, the variations in the instantaneous velocity distribution.

It is well known that at the critical the flow alternates between laminar and turbulent motion, changing abruptly from one to the other. In fact, Osborne Reynolds\* noticed two such phases; when increasing the speed from laminar, the first sign of unsteadiness is a wandering of the filament line which he used as indicator. Before finally breaking down into turbulence, however, he found that the filament line returned to its rectilinear form. The suggestion has also been made,† that, over the critical range, the velocity distribution oscillates irregularly between that appropriate to laminar flow (which in a square pipe is very nearly parabolic, see fig. 16 (b)) and that for turbulent flow. It is possible, therefore, that an inspection of the velocity distribution through the critical range by an instantaneous method such as the present might yield some further information on these points.

\* 'Phil. Trans.,' vol. 174, p. 935 (1883).

† Bond, 'Proc. Phys. Soc.,' vol. 43, p. 46 (1931).

For optical convenience the gaps were arranged in a vertical line, although it was realized that at the low speeds contemplated the convection would be appreciable near the wall and unsymmetrical owing to the gradients near roof and floor. An idea of the importance of the buoyancy could be obtained by increasing the heat energy in the spark for the condition shown in fig. 17 (c). The effect of an appreciable increase was observed to be negligible, but it was later decided to eliminate it altogether by placing the row of gaps horizontally, for reasons which will be given immediately.

To obtain the most favourable conditions, the greatest length of pipe (60 diameters) was used in conjunction with the large bell-mouth. A record of the flow just below the critical velocity is shown in fig. 17 (c); the velocity on the axis was 5.3 ft./sec. The most obvious feature of this record is the marked asymmetry of the velocity distribution. The maximum velocity occurs about midway between the axis and the roof of the pipe.

The theoretical distribution for the axial plane of a square pipe, calculated from formulæ given by Cornish\* has been plotted in fig. 16 (b). A parabola has been drawn on the same figure for comparison with the case of a circular pipe. The distribution shown in fig. 17 (c) is not even approximately parabolic, but is displaced upwards as if in consequence of the buoyancy of the spots due to their heat. This convection effect is certainly present and is important near roof and floor where it causes an over- and an under-estimation of the velocity respectively. This is in the correct sense to agree with the observed distortion of the distribution curve, but it is not sufficient to account for it entirely, as an inspection of affairs near the axis will reveal. The spot which starts on the axis is seen to rise slightly as it moves downstream. This is a simple convection effect. Presumably all spots in the same group rise by the same amount (except very near the roof) since all are created at the same instant. On account of the increasing velocity gradient near roof and floor this convection exerts a greater distortion at these points than on the axis. A spot of given velocity moving up into a region of lower velocity will therefore tend to over-estimate the speed at the point at which it has arrived when photographed, and conversely for a spot moving into a region of higher velocity. The axial spot, however, in a symmetrical flow, will move into a region of lower velocity whichever way it goes and must therefore indicate a maximum speed. Convection cannot therefore account for the fact, visible in the record, that spots *starting* half-way between the axis and roof indicate an appreciably

\* 'Proc. Roy. Soc.,' A, vol. 120, p. 691 (1928).

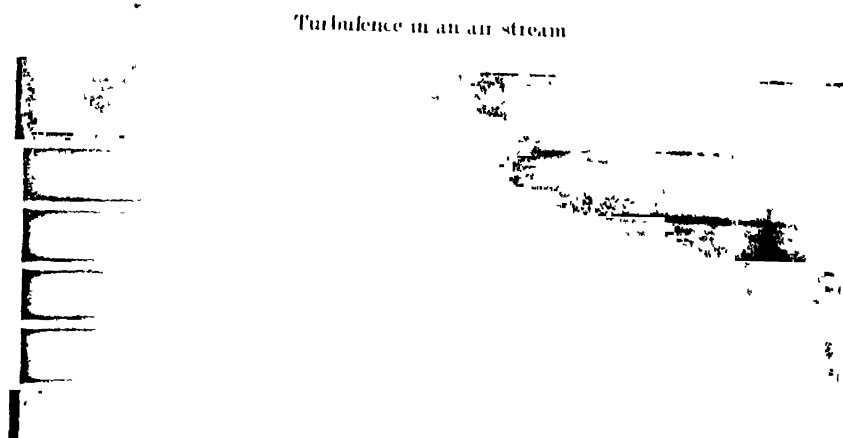


FIG. 1. Shadows of hot air filament lines produced by fine heated wires

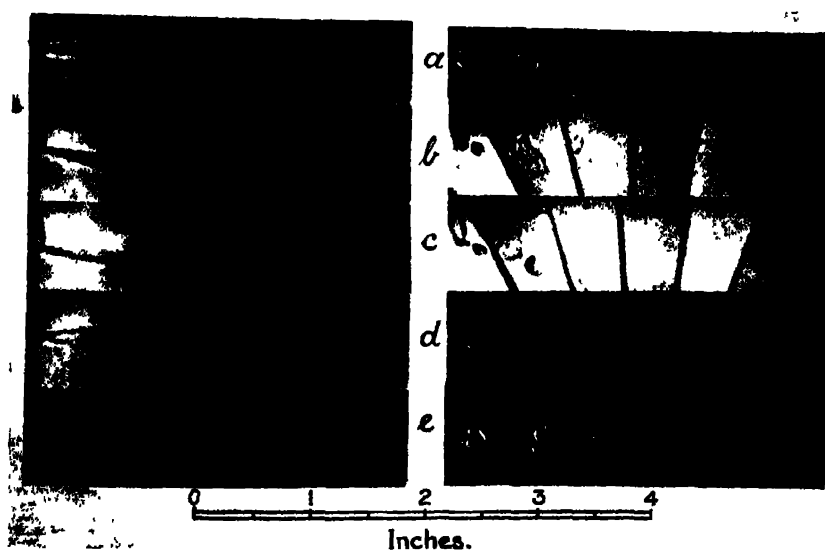


FIG. 2—Shadows of hot elements of air produced by electric sparks.

Flow on the axis of 3 inch square pipe.

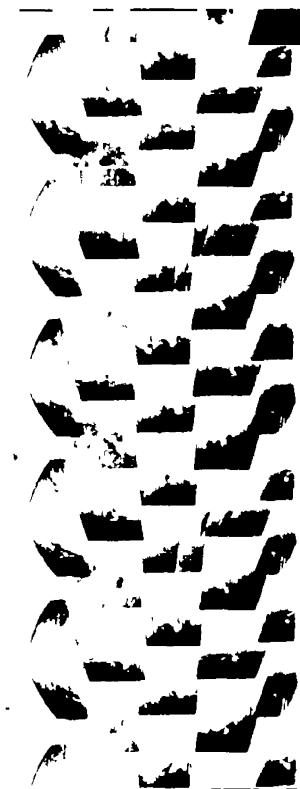


FIG. 3.  $U_c = 9.7$  ft. sec.  
Spark frequency = 26 per sec.

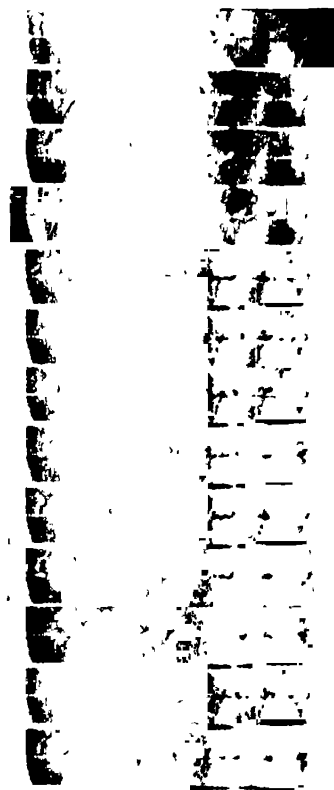


FIG. 4.  $U_c = 12$  ft. sec.  
Spark frequency = 130 per sec.  
No. of pictures per sec. = 65



FIG. 6.—Photograph of apparatus as arranged for measuring  $u$  and  $w$  only.

..... Wind direction. Vertical marks are 1·27 inches apart.

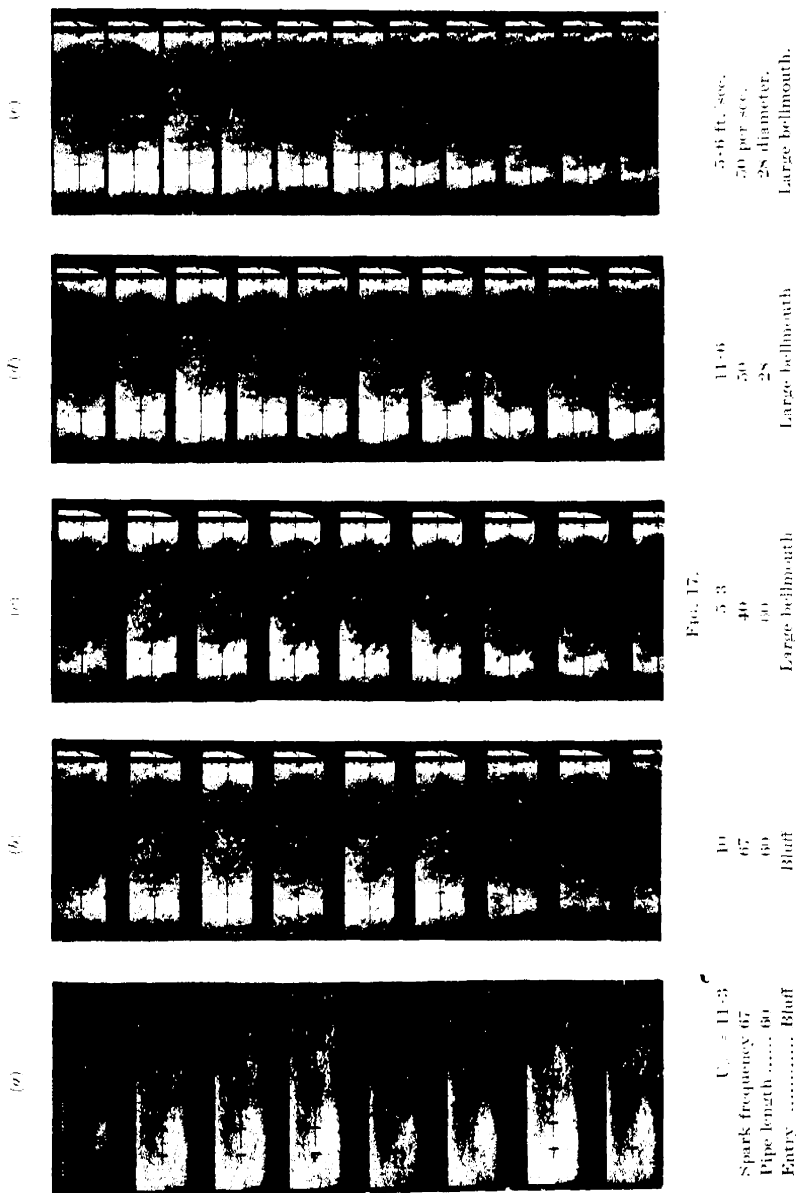


FIG. 17.





greater speed than that indicated by the axial spot. The speed must be actually greater.

On this account the spark gaps were re-arranged in a horizontal line, so that buoyancy would not carry them into regions of sensibly different velocity. The effect of this was generally, but by no means always, to induce symmetry. On several occasions unsymmetrical distributions persisted for some time. Further it was frequently noticed that the axial spot, even for symmetrical distributions, indicated a lower velocity than that on either side of it.

### *Quasi-Laminar Flow.*

On reducing the speed from the turbulent condition, in which the spots, as viewed on a screen, were oscillating erratically, there occurred a change to what may be called quasi-laminar flow. This was characterized by the sudden cessation of the oscillations of the spots. All apparent motion of the spots was not suppressed, however, as the speed of the layer represented by any one row of spots was not constant but varied slowly, often by nearly 100%, in a manner which suggested the wandering of the filament line already mentioned, particularly as a change, *e.g.*, a retardation, originating on one side of the pipe often passed slowly across it affecting each layer in turn which would then resume its earlier velocity. The rows of spots appeared to close up and then open out again in succession.

This behaviour persisted over a considerable range of velocity without any sign of the abrupt spasms of turbulence appearing. The flow, therefore, although definitely and permanently non-turbulent, was not in the ordinary sense streamline either, although the motion was practically laminar. Actually, at the moment when a change passed across from one layer to the next, it was accompanied by a very slight lateral shift of the row of spots.

Nearer the critical velocity the flow was sometimes observed for short periods to be turbulent on one side of the pipe and non-turbulent on the other, and occasionally turbulent at the centre and not turbulent at the walls.

### *Von Kármán's Theory of Turbulence.\**

A theory of turbulence is developed by von Kármán for two-dimensional motion between parallel walls. The turbulent as well as the mean motion is also treated as if it were two-dimensional, though it is known experimentally

\* 'Nachr. Ges. Wiss, Göttingen,' p. 58.

that when a steady state has been reached, disturbances originally two-dimensional eventually become three-dimensional.\* Although the theory is not strictly applicable to the flow in a square pipe its applicability is likely to be more limited by the assumption that the disturbances are two-dimensional than by the fact that the mean flow is two-dimensional. Hence it may not be unfair to test some of the conclusions of the theory in the present case, particularly as it appears, from fig. 15 of this paper, that over the central region of the pipe  $v_\sigma$  and  $w_\sigma$  are nearly equal and constant.

In von Kármán's theory the relation between  $u$  and  $v$  is independent of position. In Table VI the ratio  $u_\sigma/v_\sigma$  is given for different values of  $y/s$ .

Table VI.

$y/s$ .	$u_\sigma/v_\sigma$ .	
	$U_\sigma m/\nu = 3000$ .	$U_\sigma m/\nu = 9300$ .
0	1.17	1.17
0.4	1.65	1.72
0.6	1.50	1.85
0.8	1.51	2.20

The results show that  $u_\sigma/v_\sigma$  is greater than unity everywhere and increases considerably on moving away from the axis. Von Kármán suggests that an increase in this ratio would have been expected, but says that if it were true "it would not be possible to fix the condition by a single length† and the mean magnitude of the oscillations."

If the mean flow were made to approach the two-dimensional condition by increasing the height of the pipe, it is more likely, as argued previously (p. 204) that  $w_1$ , the component assumed zero in the theoretical case, rather than  $v_1$ , would tend to equality with  $u_1$ .

Another result of the theory is that the turbulence is zero at the centre. Though this is not supported by the results of the ultramicroscope or of the present tests, it is interesting to notice that the turbulence is a minimum on the axis and that under special conditions, before the turbulence has had time to develop completely, it may be practically zero there although turbulent everywhere else. The example shown in fig. 17 (d) illustrates this.

\* Fage, 'Aero. Res. Ctee., Rep. Mem. No. 1510' (1932).

† I.e.,  $l$ , the "mixing path."

The author would like to acknowledge the assistance of Mr. A. F. Brown B.Sc., and of Miss D. M. Yeatman in the reduction of the observations.

*Summary.*

A method has been developed of measuring the turbulent components of the velocity in a stream of air by producing a series of sparks which heat up small elements of air whose motions can then be made visible by means of the Schlieren method. Cinematograph records are made which can be analysed to obtain the mean speed at any point in the stream and the three components of the deviations from the mean velocity due to turbulence. Statistical values of the mean turbulence are then obtained and the relation between them and the maximum values of the turbulent velocity components is found.

The method has been applied to the flow of air through a pipe of square section at Reynolds' numbers of 3000 and 9300.

By using a group of seven spark gaps arranged in a row across the pipe and discharged simultaneously, the instantaneous velocity distribution is observable both above and below the critical Reynolds' number. Changes in the velocity distribution occurring in this region have been studied.

*Results.*—Measurements of the maximum values of  $u$ ,  $v$ , and  $w$ , of the turbulent velocities at various points across the pipe show good agreement with values obtained with the ultramicroscope.

The mean turbulence has been expressed in terms of the root-mean-square values,  $\sigma$ , of the components. This quantity has been used to calculate the theoretical frequency distributions given by the "error law," which depend only on this parameter, to test how far this law represents the observed distributions of  $u$ ,  $v$ , and  $w$ . The results show that the two distributions are closely similar for all positions across the pipe.

The relation between the maximum, average, and root-mean-square values of the turbulent components may be stated approximately as follows: the maximum value at any point is roughly three times the root-mean-square value, or 3.75 times the average value, neglecting sign. This also is true for all points across the pipe, and for all three components.

The distribution of turbulence was measured from the axis to a point one-tenth of the breadth of the pipe from the wall. It shows a minimum near the axis and rises at an increasing rate as the wall is approached.

There appears to be but little definite phase relation between the maximum values of the turbulent components.

*Experimental Investigation and Analysis of the Velocity Variations  
in Turbulent Flow.*

By L. F. G. SIMMONS, M.A., A.R.C.S., and C. SALTER, M.A.

(Communicated by G. I. Taylor, F.R.S.—Received December 29, 1933.)

[PLATES 4-7.]

*Introduction.*

Of the methods used to record turbulent motion in air the use of a hot-wire anemometer has received most attention; with the result the conditions under which a hot wire will accurately record the velocity variations associated with this type of motion are now well known. Many records have been obtained, all of which confirm the irregular character of turbulence, whether produced in the free stream of the wind tunnel, or in the disturbed region near the boundary of a model. Various attempts have been made from time to time to analyse the results by statistical methods, but partly owing to the difficulty of obtaining lengthy records sufficiently representative of average conditions, and partly because of the labour involved in the process of reduction, no success appears to have been achieved in this direction. Within the last year a new method of analysis suggested by Professor G. I. Taylor, which is also capable of discriminating between different types of turbulent motion, has been developed and applied to the analysis of eddying flow in a wind tunnel. The method depends on the use of a hot-wire anemometer for recording velocity variations; but in place of an Einthoven oscillograph a special type of oscillograph having a mirror attached to the suspension fibres is used so that currents produced by the velocity changes are indicated by the horizontal motion of the reflected image of an illuminated slit. If the image falls on a dropping plate in the usual way, it traces the details of the turbulent motion and provides a record in the form of a black line on a light background; that is, the reverse of that given by the Einthoven oscillograph. This all-important difference gives the new method a power of analysis that the older one lacks; for, if instead of moving across a dropping plate the reflected image moves across a stationary plate, tracing and retracing its path in accordance with the imposed changes of speed, the plate on being developed displays a dark band of variable density representing the aggregate effect of every velocity change occurring during an experiment, which may last as long as 5 or 10 minutes.

It thus constitutes a detailed record whose density at any point determined photometrically measures the total time of exposure at that point. Such information enables one to calculate the relative number of times a given displacement has been reached or exceeded in the course of an experiment. What is more important, when used in conjunction with the known relationship between the displacement of the image and wind speed, it provides sufficient data from which frequency curves can be plotted indicating the actual distribution of the recorded velocity variations.

During its development the method was applied to an examination of the flow in the wind tunnel, and some interesting facts were revealed about turbulent flow and the changes of motion brought about by the mixing and diffusion of a regular eddying system during its passage downstream.

### *Choice and Scope of Experiments.*

The experiments were conducted in a 1-foot wind tunnel, and were confined in the first instance to an examination of the motion in a system of eddies purposely introduced by a grid placed across the stream. The grid was composed of horizontal and vertical strips 0.5 inch wide, forming a mesh of 1-inch square holes; by this means velocity changes were induced which were about six or seven times larger in magnitude than those normally present in the wind tunnel, and could be traced for some distance downstream. To enable the hot-wire anemometer to record the changes accurately, most of the experiments were conducted at wind speeds of about 5 ft./sec., at which speed the main eddy frequency was of the order of 55 per second. In subsequent experiments, made without the grid for the purpose of analysing the turbulence in the free stream, the highest frequency registered was 50 per second.

### *The Hot-wire Anemometer.*

Consideration of the conditions under which a hot-wire can best be made to record velocity variations led to the use of a fine wire of 0.0025 mm. diameter in the present experiments. This was made from Wollaston wire soldered across two supports, the silver covering being removed by nitric acid leaving the platinum core exposed along a length of 0.5 mm. When used for measuring velocities a short wire such as this indicates approximately the changes occurring at a point; it has a small thermal lag, and even in an uncompensated circuit is capable of recording disturbances of a high frequency. Dryden and Kueth\*

\* The measurements of fluctuations of air speed by the hot-wire anemometer, 'Nat. Adv. Ctee. Aero., Report No. 320' (1929).

have shown that the accuracy of response is governed not only by the thermal lag but also by the temperature of the wire, which should be as low as the sensitivity of the recording arrangements permits. The effect of thermal lag on the response characteristics at 200° C. (the temperature of the wire in most of the experiments) is shown by the figures in Table I, calculated from Dryden and Kueth's formulæ for the values of  $r$ , expressing the ratios of the recorded

Table I.

Frequency.	$r$ .	$\theta$ .
		°
50	0.996	6.0
100	0.983	11.0
150	0.953	17.6
200	0.922	23.0
400	0.763	40.3

amplitude as a fraction of the true amplitude, and  $\theta$  the phase lag, where for the particular wire used at 200° C.,

$$r = \frac{1}{\sqrt{1 + 1.142 \times 10^{-7} p^2}}, \quad \theta = \tan^{-1} 3.38 \times 10^{-4} p,$$

$$p = 2\pi \times \text{frequency}.$$

It will be observed that the reduction of amplitude through thermal lag is less than 1% for the eddy frequencies met with in the experiments. The resistance of the wire at any instant therefore gives a true measure of the velocity existing at that instant, and its value can be estimated from the law of cooling, namely,

$$i^2 R = (a + b \sqrt{V}) \frac{R - R_0}{\alpha R_0}, \quad (1)$$

where

$i$  = heating current,

$R$  = resistance of the hot wire,

$R_a$  = resistance of wire at air temperature,

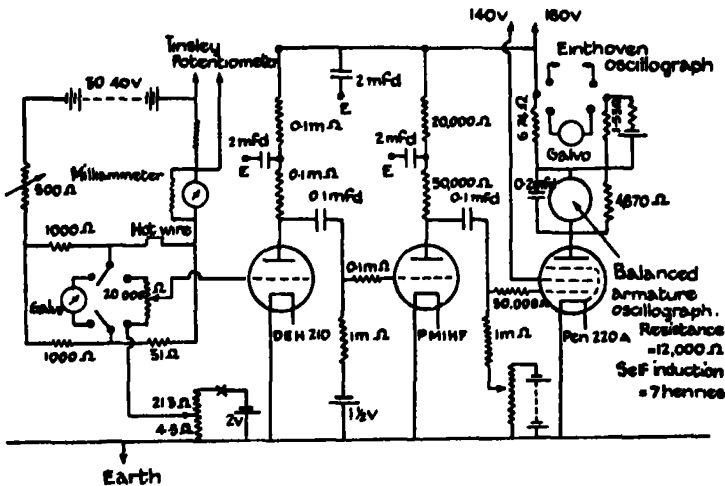
$R_0$  = resistance of wire at 0° C.,

$\alpha$  = temperature coefficient of resistance,

$V$  = wind speed.

### *The Electrical Circuit.*

To secure maximum sensitivity of the electrical circuit, it is desirable to keep the heating current through the wire as nearly constant as possible, otherwise the change of potential across it following a change of wind speed is partly neutralized by the current change. A convenient circuit is shown in fig. 1. In this the hot-wire forms one arm of a Wheatstone bridge, having an adjacent arm of 1000 ohms and other arms of 1000 and 31 ohms respectively. The heating current is supplied by accumulators, and can be adjusted by means of a variable resistance and measured by a milliammeter; or if necessary it can be more accurately determined from potentiometer measurements across a standard resistance of 29 ohms, connected in series with the battery. The former method is usefully employed as a guide in adjusting the current



**FIG. 1.—Wheatstone bridge circuit and amplifier.**

required to balance the bridge at any given speed ; but the latter method is essential for the accurate determination of the constants  $a$  and  $b$  in equation (1).

When the bridge is used for recording fluctuating velocities, the current is adjusted so that the average potential across the bridge, as indicated by a slow-period galvanometer, is zero. Such balance is made possible by the inability of the galvanometer to follow the potential changes.

In the experiments with a heating current through the wire of 20 milliamperes the maximum potentials produced across the bridge by the cooling effect of the air varied from 0.01 volts for free stream turbulence to 0.1 volt for large disturbances. Potential changes of this order of magnitude could not be recorded unless first amplified by a valve circuit, a description of which follows.



*The Amplifier.*

An amplifier of the audio-frequency type, with two stages of resistance capacity coupling, possesses certain advantages over a direct-current amplifier or one with transformer coupling, since it is free from leakage troubles common to the former and reproduces more accurately than the latter any irregular variation of potential. With 0.1 microfarad coupling condensers and grid leaks of 1 megohm, the amplification is found to be practically constant over a range of frequencies from 5 to 2000 cycles per second; but below 5 per second the response curve gradually falls to zero. Consequently, an amplifier of this kind takes no account of slow variations of wind speed due to the heating of the tunnel motor or those produced by a swirling motion about the axis of the wind tunnel. Thus it conveniently ignores any slow disturbances and only responds to those which, in the present experiments, are legitimately connected with the eddying system or with turbulence proper.

On reference to the diagrammatic sketch of the amplifier in fig. 1, it is seen that the first two valves are decoupled in the usual manner by resistances and condensers of 2 microfarad capacity—although the use of fully charged accumulators on the high tension supply reduces the risk of any back-coupling and so makes the need for this precautionary measure somewhat redundant. Some care is required in the selection of the valves: a suitable arrangement consists of a D.E.H. 210 high impedance valve, followed by a P.M.I.H.F. of medium impedance, and a pentode Pen 220A. The grid of the first valve is biased to about 0.4 volt; the second to  $1\frac{1}{2}$  volts; and the third to a suitable voltage, depending on the loading, by means of an adjustable potentiometer. The pentode is chosen because of its high mutual conductance. Its screen should be maintained at 140 volts and not decoupled like the anode circuits of the other valves. Any attempt to do this, using a condenser as small as 2 mfd., leads to a reduction of the amplification of the lower frequencies by about 30%, because under such conditions the *time constant* of the screen circuit is not large enough to prevent the screen voltage from varying in sympathy with the changes impressed on the grid. Special care is necessary to shield the amplifier from extraneous disturbances of high frequency. A satisfactory method of screening consists of encasing the amplifier in a metal box, with metal partitions between the stages, to reduce any possible interaction between them to a minimum. Connecting leads to the first valve are inserted in a metal tube attached to the box containing the amplifier, all parts of the screening system being connected to a common earth.

In order to prevent overloading the amplifier, provision must be made for controlling the input. An effective method is to connect the grid of the first valve to a tapped potentiometer shunted across the bridge as shown in fig. 1. By selecting the tapping point, any proportion of the out-of-balance potential can, in this way, be amplified at will.

### *Oscillographs.*

For recording one or other of two oscillographs can be put in the anode circuit of the pentode: the first, an Einthoven oscillograph, is convenient for indicating the wave form; the second, described by Wynn-Williams and Ward,\* hereafter referred to as the balanced armature oscillograph, can be used for analysing the variations of e.m.f. applied to the first valve. With the former, since current changes are indicated by the motion of a shadow image of the fibre across an illuminated background, it is impossible to secure on a stationary photographic plate a record covering a long period of time, such as can be obtained with the second instrument. This is based on the design of a loudspeaker movement incorporating a balanced armature; it consists of a coil surrounding a short rod of soft iron attached at the ends to a bifilar suspension, the tension in which can be adjusted to give a high natural frequency. A beam of light reflected from a mirror fixed at one end of the rod gives a measure of the deflection. If care is taken to ensure that the iron armature is perpendicular to the resultant field produced by the combination of the main field and the field induced by the D.C. component of the current through the coil, the response to current variations is linear over a considerable range. This precaution is especially important if, as in the experiments, the oscillograph is connected to the last valve, so that a large proportion of the anode current (amounting to about 5 milliamperes) passes through the coil.

The moving parts and coil of the instrument, as originally designed and constructed at the Cavendish Laboratory, were immersed in a bath of oil. At low frequencies the damping obtained is normally sufficient; but in order to make the instrument dead-beat under the conditions of the experiments, it was found advisable to shunt it with a condenser of 0.2 microfarad capacity. Fig. 6, Plate 4, shows the result of a small e.m.f. suddenly applied to the instrument: it will be observed that the overshoot is small. A similar record indicating the satisfactory response characteristics of the Einthoven oscillograph is given in fig. 7, Plate 5.

\* 'Proc. Roy. Soc.,' A, vol. 131, p. 391 (1931).

*Calibration of Amplifier and Oscillographs.*

In any prolonged investigation a frequent check has to be kept on the calibration of the amplifier. In the method employed a small sinusoidal variation of voltage is applied to the first valve, and the extreme excursions of the beam of light or shadowgraph recording the current change are recorded photographically. The e.m.f. is generated by a small coil rotated in the magnetic field produced by two large Helmholtz coils which are energized by a constant current. Various precautions are taken to secure a constant and uniform field. Among these may be mentioned the metal box used for screening the coils from stray fields and other disturbances; also the long, phosphor bronze rod, to which the moving coil is attached, by which it is rotated by a motor so far removed from the coils that it has no distorting influence on the field.

Calibrations were made at frequencies between 10 and 20 per second; the e.m.f. of the generator at any frequency being estimated from vacuum-junction measurements of current passing through a known resistance. Lastly each oscillograph was calibrated with small direct currents to define the range over which a linear relationship holds between the current and the deflection.

*Distortion.*

As any departure from the linear law introduces distortion in the wave form of the output, the maximum current variations were always kept well within the capacity of the apparatus. Care was also taken to set the grid bias to an appropriate value for each valve, in order to restrict the grid swing to the linear range of the grid-current curve; but during the early stages of the experiments a certain amount of distortion arose from the overloading of the last valve, as was clearly indicated in the records by the addition of a second harmonic superimposed on the fundamental. Later the use of a pentode in place of a power valve in the last stage effectively removed this cause of uncertainty. Two other sources of distortion, however, remained, both having their origin in the balanced armature oscillograph; one due to the use of the condenser for damping, the other to hysteresis in the small iron armature. The first reduced the amplification factor and introduced a phase change, whose magnitude could be calculated from the known resistance of the oscillograph and equivalent shunt resistance of the condenser. Both the amplification and phase varied with frequency, and thus the distortion had a selective effect. At a frequency of 100 per second the calculated reduction in amplitude was 6.9% and the phase change 40°; but at the eddy

frequency the corresponding figures were 1% and 20°; the distortion effects were therefore small. The second source of distortion led to the presence of a third harmonic of small amplitude in the output, thereby influencing the wave form directly, though only by a small amount.

Fig. 5, Plate 4, represents a typical calibration record, taken on a plate falling with an accelerated motion, and illustrates the output wave form given by the balanced armature oscillograph from a sinusoidal e.m.f. applied to the amplifier. Analysis shows that, after allowance is made for the acceleration of the plate, if the applied voltage is denoted by  $E \sin \omega t$  the output current at a frequency of 100 per second is of the form

$$KE (\sin \omega t + 0.03 \sin 2\omega t + 0.03 \sin 3\omega t)$$

approximately, where  $K$  is a constant of the amplifier and recording circuit.

### Velocity Scales of Records.

In the absence of a direct method of calibration, the velocity scale for the oscillograph records were calculated from a knowledge of the electrical constants including those of the amplifier and the oscillograph and the law of cooling of the wire in the following manner.

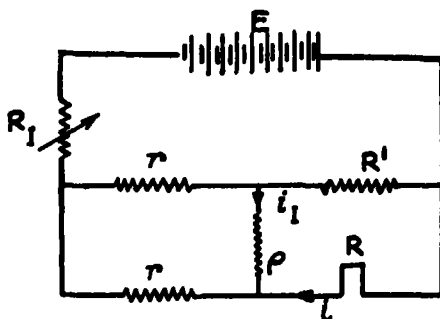


FIG. 1A.

Let

$R$  = the resistance of the hot wire at a velocity  $V$ .

$R'$  and  $r$  = respective resistances of the remaining arms of the bridge.

$i$  = current in the hot wire.

$i_1$  = out-of-balance current.

$R_1$  = external resistance.

$E$  = e.m.f. of the battery.

$\rho$  = resistance of the potentiometer across the bridge.

Then to determine the out-of-balance potential produced by a deviation  $V - V_0$  from the mean speed  $V_0$ , we require expressions for  $i$  and  $i_1$  in terms of the various constants. These can easily be shown to be

$$\frac{i}{E} = \frac{r\rho + (2r + \rho)R'}{\epsilon}, \quad (2)$$

and

$$\frac{i_1}{E} = \frac{(R - R')r}{\epsilon}, \quad (3)$$

where

$$\begin{aligned} \epsilon = (R - R') \{ (2R_1 + R' + r)(r + \rho) + (rR' - \rho R_1) \} \\ + (2R_1 + R' + r) \{ R'(2r + \rho) + r\rho \}. \end{aligned}$$

Also from (1)

$$i^2 R = (a + b\sqrt{V}) \frac{R - R_0}{\alpha R_0}. \quad (4)$$

If the velocity changes are small

$$\frac{dV}{di_1} = \frac{2\alpha\rho R_0 i^2 \sqrt{V_0}}{b(R - R_0)^2 E} \left( \frac{2R^2 R_1}{r} + 2R_0 R_1 + 2R^2 + R_0 r \right) \text{ approximately; } (5)$$

an equation that can be used to calculate the small change of velocity corresponding to a change of potential  $\rho di_1$ .

More generally, for larger variations of velocity, it is necessary to establish a relation between  $i_1$  and  $V$  by solving equations (2), (3), and (4). This relation is best shown graphically, and can be obtained by assuming different values for  $R$  and calculating the corresponding values of  $i_1$  and  $V$  by means of the equations. Then from a knowledge of the deflection of the oscillograph produced by a given e.m.f. applied to the amplifier, a calibration curve can be plotted, as in fig. 3.

### *The Optical System.*

Satisfactory records for statistical analysis can be secured only if care is taken to provide a constant source of light giving a narrow band of approximately uniform distribution for the reflected image. An opal lamp run at a constant current serves admirably for the purpose; with a current somewhat lower than the normal rating, the life of the lamp is considerably increased, and the intensity of the light is found to remain constant over a long period.

On reference to fig. 2, showing diagrammatically the details of the optical system used in conjunction with the balanced armature oscillograph, it will be seen that the virtual source is an illuminated slit, the emergent beam from which is reflected by the mirror to form an image about 1.2 mm. wide on the

photographic plate. A small-angled prism compensates for the dispersion caused by the damping liquid in the oscillograph; and a small convex lens, also in the reflected beam, corrects the effect due to distortion of the mirror, such as was found to occur in the experiments and was attributed to the dissolving action of the liquid on the wax employed for fixing the mirror to the iron armature.

The experiments are necessarily conducted in a dark room, wherein stray light can be reduced to a minimum. Light reflected from the surfaces of the various components forming the optical system can be suitably deflected so as not to fall on the plate. If slow process plates are used successful records can be obtained with exposures of 4 to 5 minutes without the general fogging exceeding a density greater than 1 or 2% of the maximum recorded.

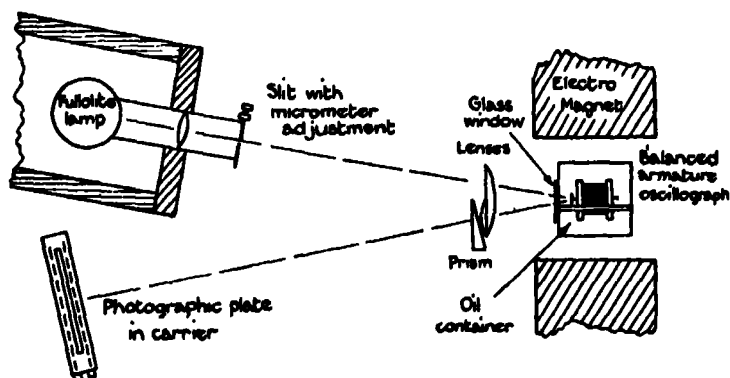


FIG. 2.—The optical system.

### *Method of Recording.*

A standard procedure is adopted for recording, developing, and fixing the negatives; in addition, each plate is calibrated separately, to allow for slight variations in developing, etc. The method employed throughout the experiments consisted, of first exposing the plate to the fluctuating band of light, moving in unison with the velocity variations for about 5 minutes; and afterwards exposing other parts of it, each for a period of 10 seconds, to the reflected beam from the mirror maintained in its position of rest. The average density  $D_{10}$  of these stationary images (usually four in number) gives a measure of the photographic quality of any plate, and thus serves as a convenient basis of reference. All density measurements of the records are therefore expressed as a fraction of  $D_{10}$ .

Fig. 8, Plate 5, is an example of a complete record made in the manner described, showing in order, upwards from the lowest, the relative distributions

of density produced by (1) velocity disturbances, (a) for exposure of 4 minutes, (b) for exposure of 5 minutes; (2) four stationary calibration bands; (3) a simple harmonic motion. The purpose of (3) will be described later.

### *Measurement of Density.*

Analyses of the records ultimately rest on the actual density distribution along them, measured with a Moll photometer. In this instrument light from a slit, illuminated by a small standard lamp, is brought to a focus in the plane of the negative, and thence focussed on a thermopile. When a negative is interposed the heat radiation received by the thermopile, and indicated by a reflecting galvanometer, varies with the intensity of the emergent beam as affected by the opacity of the negative. Thus the deflection of the galvanometer being proportional to the intensity of the heat falling on the thermopile and therefore to the amount of light transmitted, measures the photometric density. To determine the density distribution, the plate is slowly traversed across the beam of light, and the deflection of the galvanometer is photographically recorded on paper attached to a drum rotated at a uniform speed. Reference lines are added; firstly, by excluding the light from the thermopile, and secondly, by exposing the plate on each side of the record to measure the density of the general background produced by stray light. The distance between the two lines is proportional to  $I_0$ , the intensity of the incident light, whilst the ordinate of any point on a record measured from the first or total blackening line, is proportional to the intensity of the light  $I$  transmitted through the plate. The density  $D$  at any point can thus be computed readily from the ratio of two ordinates, in accordance with the formula

$$D = \log_{10} \frac{I_0}{I}. \quad (6)$$

Inspection of some photometric records obtained in the course of the investigations and reproduced in figs. 13-16, Plate 7, clearly shows the upper lines made by shielding the thermopile from the light, and the lower fog line as an extended base of each diagram. Special interest attaches to fig. 13 as representing the type of calibration record for each plate from which values of  $D_{10}$  are calculated.

### *Density-time Relation.*

So far a description has been given of the method adopted for making the records and measuring the density of the negatives. In the next stage of the analysis the time of exposure for each point on the record is deduced from the

density measurements, the conversion being made by means of the relation existing between density and time of exposure.

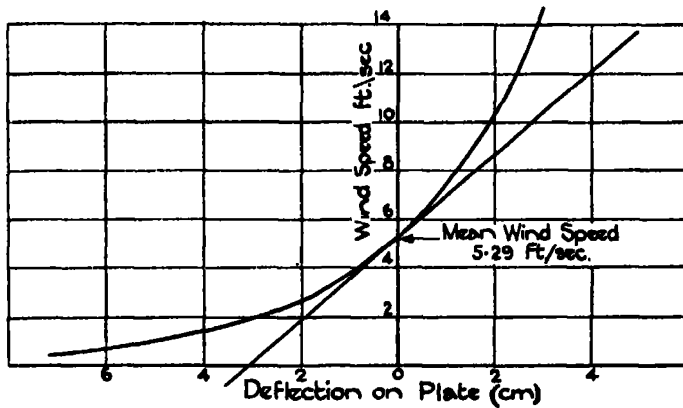


FIG. 3.—Typical calibration curve of wire and apparatus.

For continuous light of constant intensity this relationship is readily derived from records of different exposures. The results so obtained are plotted in fig. 4, where it is seen that the density is a non-linear function of the time of

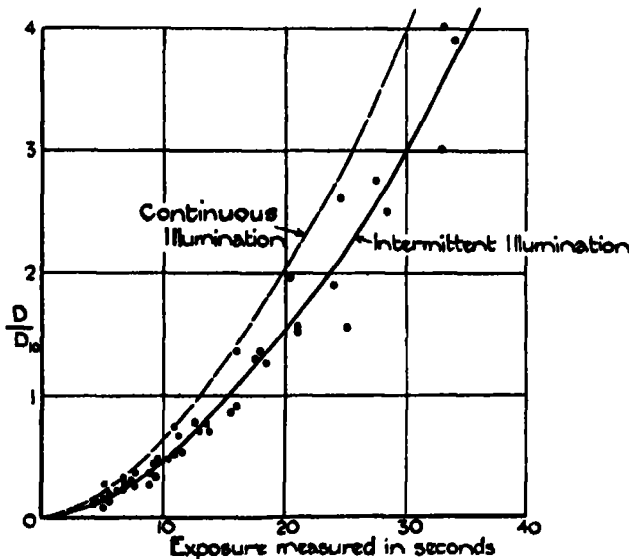


FIG. 4.—Calibration of negatives.

exposure, increasing at a greater rate as the time is increased. Owing to the failure of the plate to integrate completely the effects of discontinuous illumina-



tion, a different law holds when the light is intermittent. Having regard to the conditions of the experiments, a fair approximation to the relation required is obtained from records made with the band of light performing a simple harmonic motion. One such record is shown in fig. 8, Plate 5, and the corresponding photometer record in fig. 14, Plate 7. From the latter curve coupled with the calibration factor of the plate,  $D_{10}$ , the values of  $(D/D_{10})$  can be calculated for all displacements from the mean. Then, from a knowledge of the width of the band and the law of motion, the total time that any part of the plate is exposed can be calculated and the results plotted as  $D/D_{10}$  against  $t$ , as in fig. 4, where a mean curve is drawn to pass through the C.G. of groups of 10 adjacent observations.

#### *Stationary Plate Records.*

Stationary photographs of the speed variations were taken for a number of positions of the wire along two lines parallel to the axis of the tunnel, one passing through a strip, the other through the centre of a neighbouring hole in the grid, near the tunnel axis as indicated in fig. 25; for purposes of comparison, a similar record was made of the free stream turbulence at the same wind speed, namely, 5.3 feet per second. Most of the records were made with a total exposure of 5 minutes, and therefore represent the changes occurring over a reasonably long interval of time. In addition to the velocity record each negative, as explained, included four images for the measurement of  $D_{10}$ , but some like the specimen in fig. 8, Plate 5, also contained a record of a simple harmonic oscillation from which the density-time relation was derived. Generally speaking, each velocity record appeared as a dark band at the centre which faded gradually in tone into the light background on each side. Except in the records taken near the grid, wherein the distribution was slightly less symmetrical, the density distribution was usually found to resemble the curves of figs. 15 and 16, Plate 7, in which the density is seen to reach a maximum at the centre. Thus it follows that, since throughout the major part of the time occupied by an experiment the image remained near the middle of the record, the deviations of velocity from the mean were, in general, small.

The following analysis applied to the photometer records enables curves to be plotted exhibiting the velocity frequency of the disturbances, and also provides a method of determining the magnitude of the root-mean-square velocity.

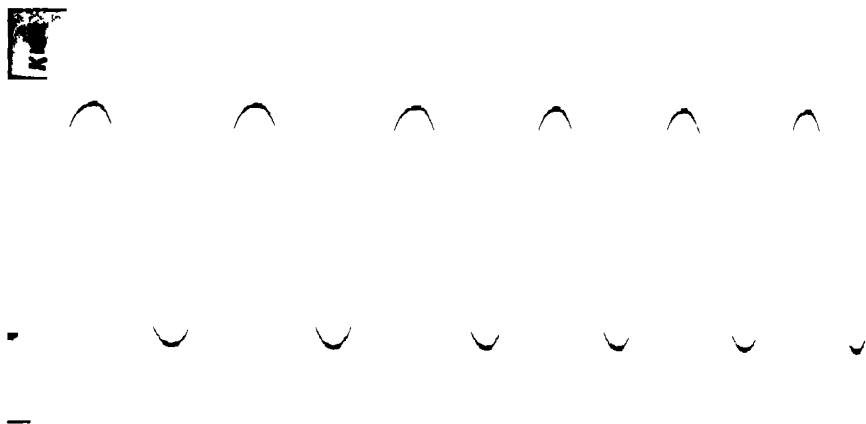


FIG. 5. Output wave form produced by a sinusoidal e.m.f.

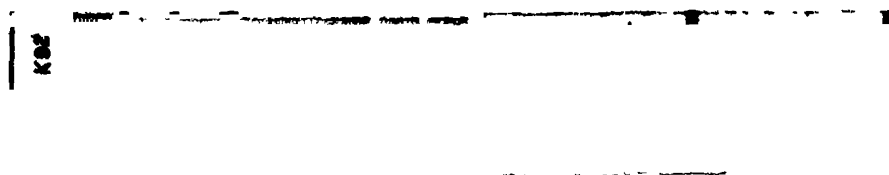


FIG. 6. Damping record of balanced-armature oscillograph.

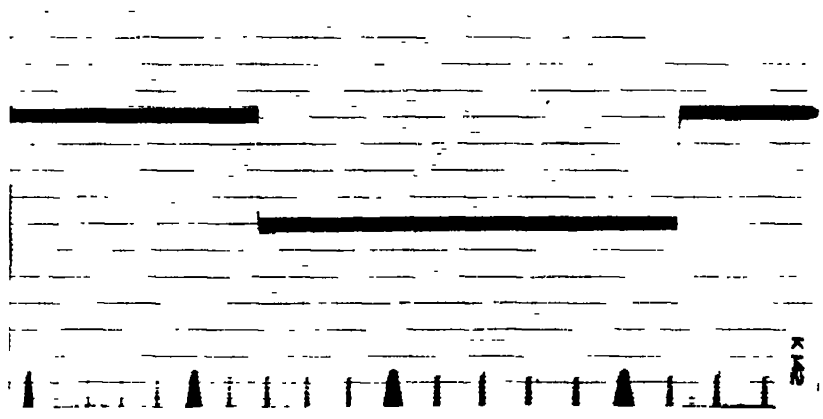


FIG. 7. Damping record of Einthoven oscillograph.

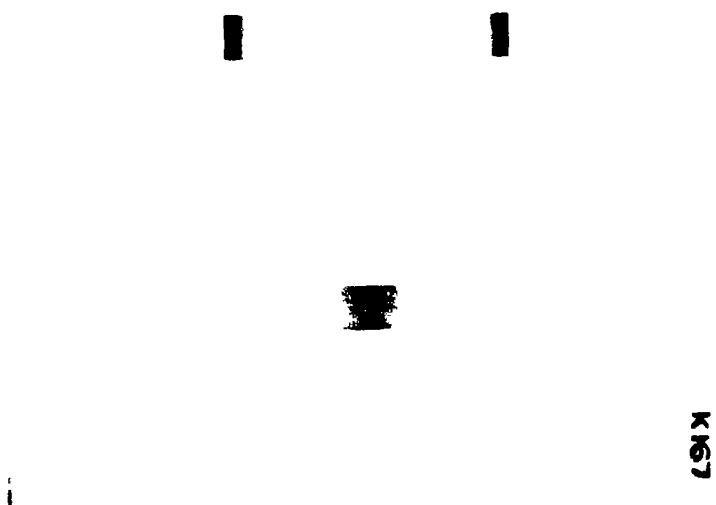


FIG. 8. Calibration record and record of disturbances at 15 inches for exposures of 5 and 4 minutes respectively.

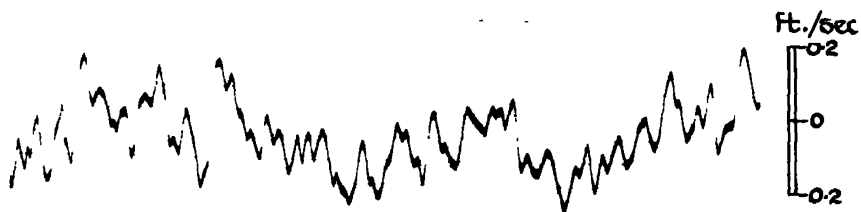


FIG.9 Free stream; mean wind speed 5.29 ft./sec.

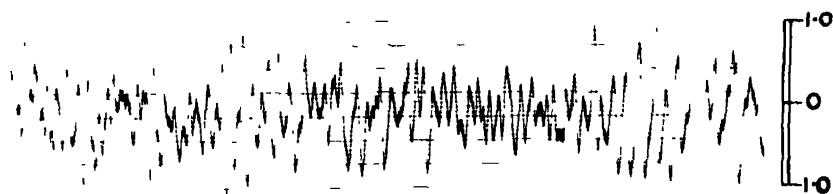


FIG.10 2 ins. downstream of B; local mean wind speed 15.3 ft./sec.

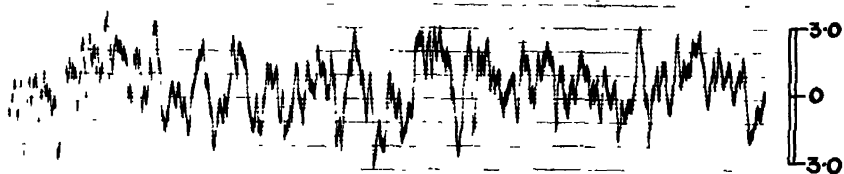


FIG.11 6 ins. downstream of A; local mean wind speed 5.19 ft./sec.

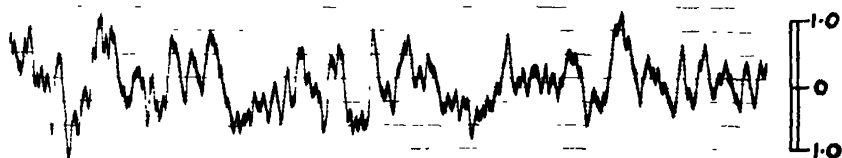


FIG.12 15 ins. downstream; mean wind speed 5.29 ft./sec.



Fig 13 Stationary images, exposure in each case 10 seconds



Fig 14 Calibration with sinusoidal osc<sup>n</sup> exposure 6 minutes

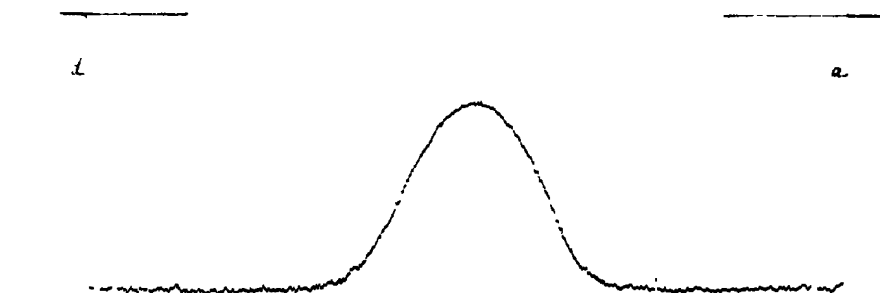


Fig 15 Free stream turbulence, exposure 3 minutes



Fig 16 Record taken 15 inches downstream of grid, exposure 5 minutes

# Analysis.

The different stages in the analysis are best explained with the aid of diagrams. Suppose the curve in fig. 17 represents a photometer record of a negative exposed for T minutes, AB being the fog line and CD, EF the total

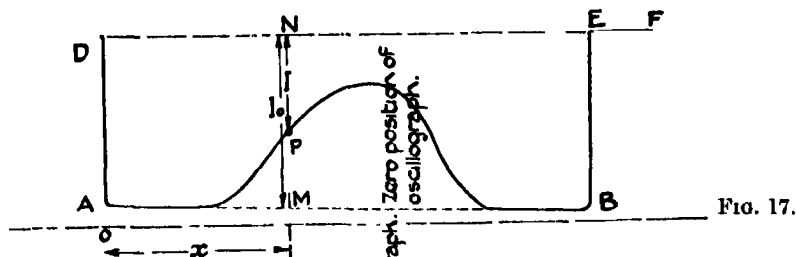


FIG. 17.

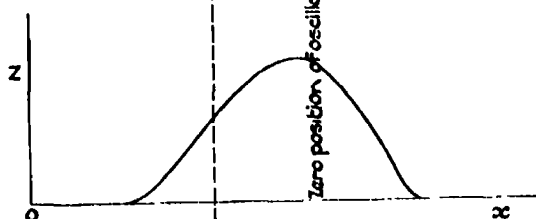


FIG. 18.

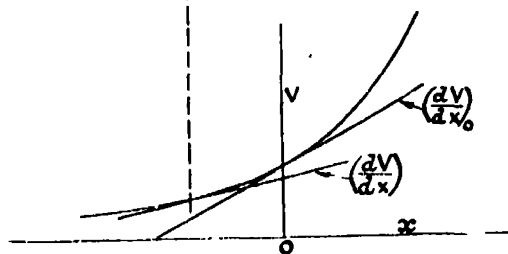


FIG. 19.

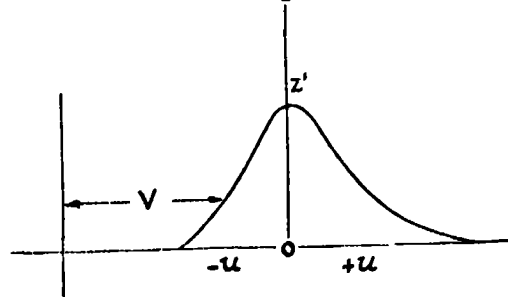


FIG. 20.

FIGS. 17-20.—Diagram showing method of reduction.

extinction line. Let P be any point on the curve at a distance  $x$  measured from some arbitrary origin O. Then the density D at P is given by

$$D = \log_{10} \frac{I_0}{I} = \log_{10} \frac{MN}{NP}. \quad (7)$$

Also as  $D_{10}$  can be measured similarly from the calibration bands, the value of  $D/D_{10}$  can be readily calculated, and the equivalent time of exposure  $t$  can be read off the calibration curve for intermittent light, fig. 4. At this stage it is convenient to plot  $z$  against  $x$ , where  $z = t/T$  as in fig. 18. It should be remembered that  $x$  is a distance measured parallel to the base of the photometer record, and as such represents wind speed in the manner shown by the curve of fig. 19, expressed by the functional relation  $V = f(x)$ ; further  $z\delta x$  is proportional to the frequency of the deflections of the image occurring between  $x$  and  $x + \delta x$ . Hence to obtain a curve of  $z'$  against velocity, where  $z'\delta V$  is to be proportional to the frequency of the disturbances between  $V$  and  $V + \delta V$ , we must transform the variables in accordance with the equations

$$V = f(x), \quad (8)$$

$$z' = z \left( \frac{dV}{dx} \right)_0 \bigg/ \left( \frac{dV}{dx} \right), \quad (9)$$

where  $(dV/dx)_0$  is the slope of the  $(V, x)$  curve at the position corresponding with the mean speed  $V_0$ , see fig. 19; the abscissa of the mean velocity  $V_0$  should incidentally coincide with the vertical through the C.G. of the area enclosed by the new curve.

If now  $F(u)\delta u$  is the frequency of velocities occurring between  $u$  and  $u + \delta u$  (where  $u = V - V_0$ , the deviation from the mean), then

$$F(u)\delta u = z' \delta V \bigg/ \int_{-\infty}^{\infty} z' dV,$$

or, since

$$\delta u = \delta V$$

$$F(u) = z' \bigg/ \int_{-\infty}^{\infty} z' dV. \quad (10)$$

Thus by dividing each ordinate of the  $(z' - V)$  curve by the area enclosed, we obtain the frequency distribution curve of velocities. In addition, curves of  $u^2 F(u)$  against  $u$  can be plotted and by integration the mean square velocity  $\bar{u}^2$  calculated from the relation

$$\bar{u}^2 = \int_{-\infty}^{+\infty} u^2 F(u) du. \quad (11)$$

### *Results of Analysis.*

The results of the analysis are summarized in the table and graphs appended, which give particulars such as velocity distribution, range of variation of  $u$ , maximum values of  $F(u)$ , and values of

$$\bar{u}^2, \quad h, \quad \frac{1}{2h^2}, \quad \text{and} \quad \frac{\int z dx}{\delta}.$$

*Velocity Distribution.*

Measurements of average velocity also reveal some interesting features of the flow downstream of the grid, particularly in relation to the mixing of the high velocity stream through the mesh and the eddying wakes in the rear of the strips. Through the interchange of momentum, the wake streams are accelerated and the high velocity jets retarded. The process continues for some distance downstream until uniform conditions are attained when the velocity reaches the free stream value. Reference to the results given in fig. 25 show, as before, that mixing extends for a distance of 10 inches at a wind speed of 5.3 feet per second; other results, not included, indicate that the distance increases with increase of wind speed.

*Values of  $F(u)$  and  $\overline{u^2}$ .*

Values of  $F(u)$  and  $u^2Fu$  plotted in figs. 21 and 22 for two positions, one at 2 inches and the other at 15 inches, downstream of the grid, present in a graphical form the frequency and energy distributions of the eddies at the two selected positions. In the wake of a strip near the grid both distributions indicate the presence of disturbances causing systematic variations of velocity in one direction from the mean, which apparently persist for some distance back. At still greater distances under the influence of diffusion the flow changes in character; and finally, when mixing is complete, the distribution assumes a symmetrical form showing that the velocity variations are equally distributed about the mean.

Further evidence of a similar nature is provided by the two series of curves in fig. 23 indicating the frequency distribution at various positions along the two lines through A and B parallel to the tunnel axis (see fig. 25). Those of the first series, plotted from records taken at 2 and 3 inches from the grid, are unsymmetrical because of the persistence of disturbances of low velocity. Farther away as the streams become more effectively mixed, the maximum amplitude is reduced and the distribution becomes more symmetrical. Finally, at, and beyond, a distance of about 10 inches, where the eddying motion is completely diffused, the curves are symmetrical and conform closely to the normal error law expressed by the relation

$$F(u) = \frac{h}{\sqrt{\pi}} e^{-h^2 u^2}. \quad (12)$$

This law also applies to the distribution of turbulence in the free stream. Hence along the line through B near the grid, where the conditions of flow resemble



those in the free stream, the velocity frequency distribution also follows the same law. As the eddying wake spreads the small disturbances increase in magnitude and reach a maximum value at some position between 3 and 6

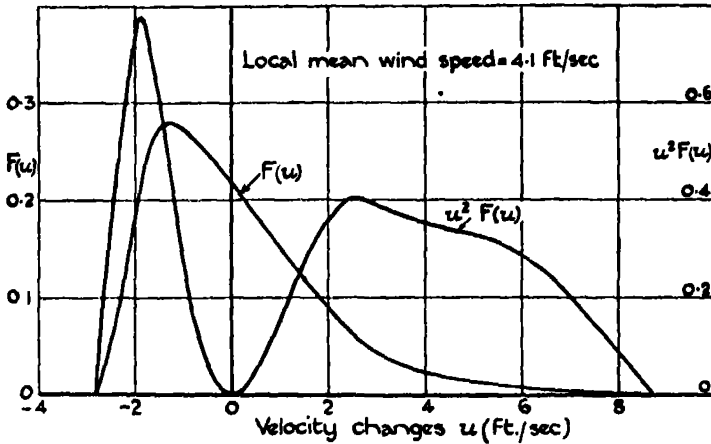


FIG. 21.

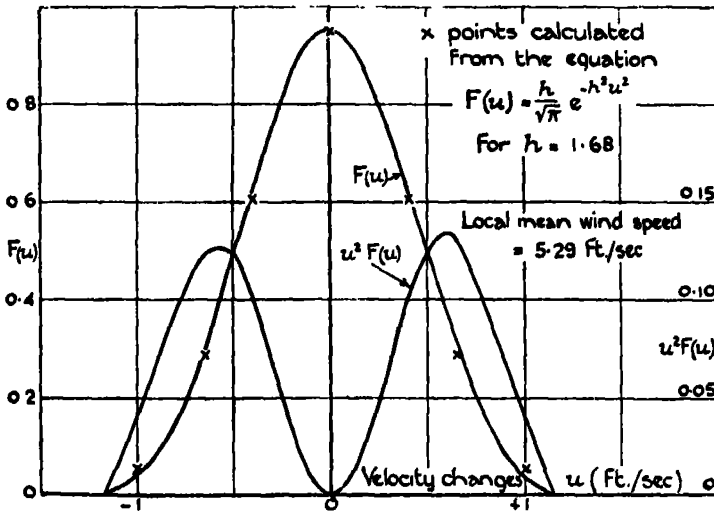


FIG. 22.

FIGS. 21, 22.—Curves of  $F(u)$  and  $u^2 F(u)$  at two positions downstream of the grid, deduced from observations taken at wind speed = 5.29 ft./sec. Fig. 21, 2 inches downstream on a line through A. Fig. 22, 15 inches downstream.

inches back, fig. 26; afterwards they become progressively smaller as the distance increases. Throughout the changing process the curves of distribution remain nearly symmetrical; the small departure from the law given above being most noticeable within a distance of 9 inches from the grid.

The curves in fig. 24 plotted on the same scale convey a clearer idea of the functional relation between  $F(u)$  and  $u$ , established from records taken downstream of the grid and also in the free stream. As explained, the curves obey

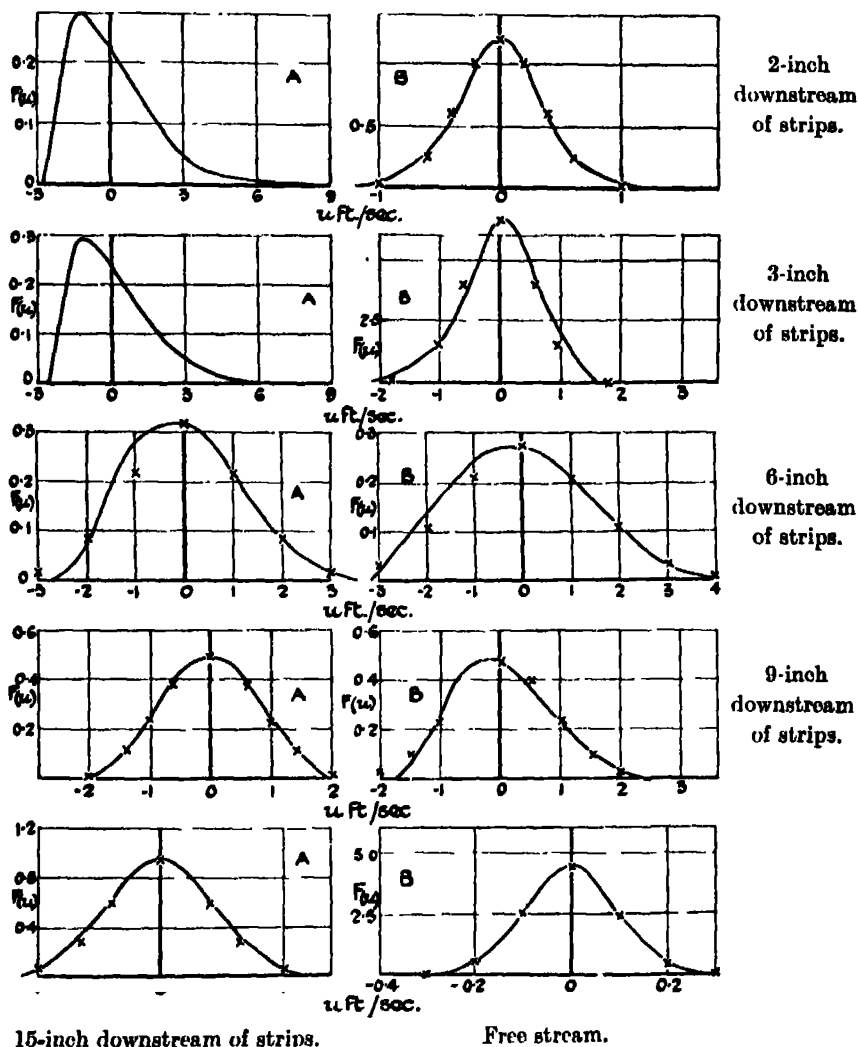


FIG. 23.—Relations between  $F(u)$  and  $u$  measured at different positions downstream of the grid and also in the free stream. Wind speed = 5.29 ft./sec. Results deduced from experiments shown by full lines;  $\times$  calculated from equation

$$F(u) = \frac{h}{\sqrt{\pi}} e^{-h^2 u^2}.$$

equation (12),  $h$  being different for each curve. The values of  $h$  are calculated from the maximum values of  $F(u)$ ; and since the mean square velocity for

the normal error law of distribution is equal to  $1/2h^2$  as shown below, the good agreement observed between the values of  $\bar{u}^2$  and  $1/2h^2$  in Table II, is of particular interest.

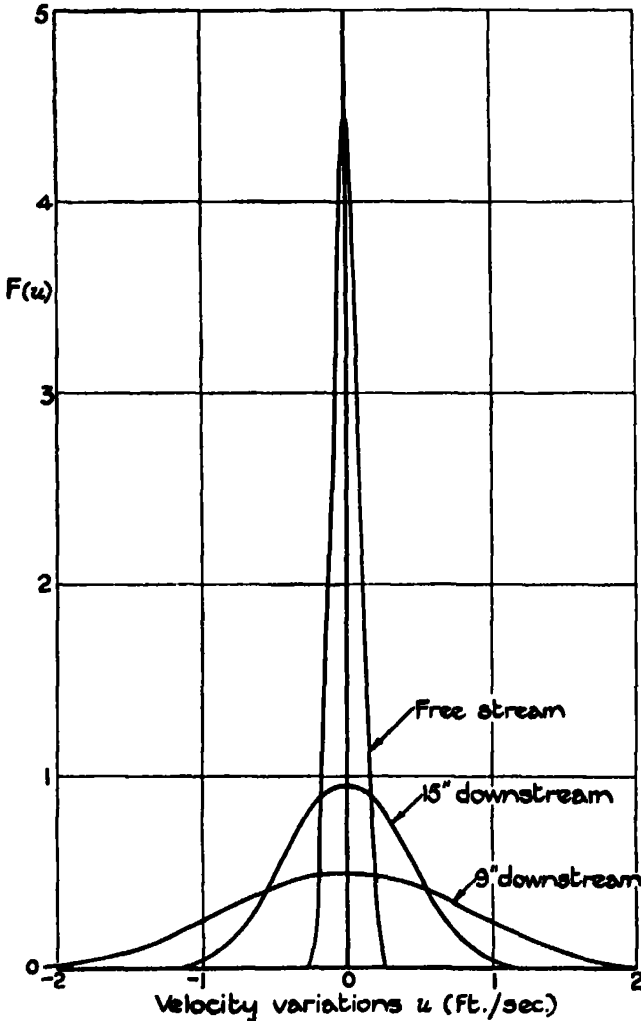


FIG. 24.—Velocity frequency curve for two positions downstream of the grid, and in the free stream. Wind speed 5.3 ft./sec.

To obtain an expression for  $\bar{u}^2$  when the frequency distribution follows the normal error law, we have

$$\int_{-\infty}^{+\infty} F(u) du = \int_{-\infty}^{\infty} \frac{h}{\sqrt{\pi}} e^{-h^2 u^2} du = 1.$$

Differentiate with respect to  $h$ , then

$$\int_{-\infty}^{\infty} \frac{e^{-h^2 u^2}}{\sqrt{\pi}} du - \int_{-\infty}^{\infty} \frac{2h^2 u^2}{\sqrt{\pi}} e^{-h^2 u^2} du = 0.$$

But

$$\int_{-\infty}^{\infty} \frac{e^{-h^2 u^2}}{\sqrt{\pi}} du = 1,$$

hence

$$2h \int_{-\infty}^{\infty} u^2 F(u) du = \frac{1}{h},$$

therefore

$$\overline{u^2} = \frac{1}{2h^2}.$$

The figures in the last column of Table II calculated from  $\int z dx / \delta$ , where  $\delta$  = width of the image, serve as a check on the accuracy of the calibration. These should be constant, namely, 60 seconds; they are, however, usually somewhat greater than this value, probably because the calibration curve of the negative, though derived from simple harmonic variations, does not completely represent the effect of the image coming to rest on reversal of its direction.

Details of the energy exchange between the wake and the jet stream are well brought out by the figures in columns 4, 7, and 8, and also shown by curves in figs. 26 and 27. Immediately in the wake of a strip the results are influenced to some extent by the large variations of direction of the flow present in that region. Even if these particular results are disregarded, there is sufficient evidence to show that the wake stream disturbances decrease in intensity continuously downstream of the grid; whereas those in the jet stream increase rapidly for the first 5 inches and then decay rapidly during the next 5 inches; at 10 inches where the eddying motion is completely diffused, they attain the same order of magnitude as those in the wake.

The maximum values of  $u$  measured along the lines through A and B, fig. 26, also exhibit the interchange taking place between the faster and slower moving air; while the coincidence of the two curves at and beyond 10 inches is additional evidence of the completion of the mixing process. Einthoven records disclose further details of the velocity changes caused by the eddy system. Those reproduced in figs. 10-12, Plate 6, enable the changes to be followed downstream, and clearly show the transition from the quasi-periodic motion near the grid to the irregular type of flow after the mixing process is completed. It should be added that experiments made under steadier wind conditions show that the initial irregularities observed in the eddying wake are due to the influence of turbulence in the free stream, fig. 9, Plate 6, which

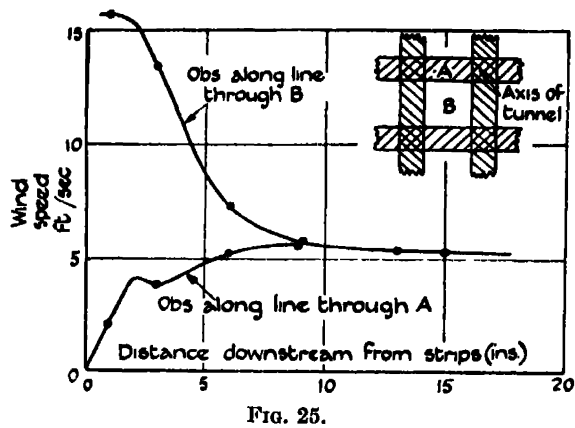


FIG. 25.

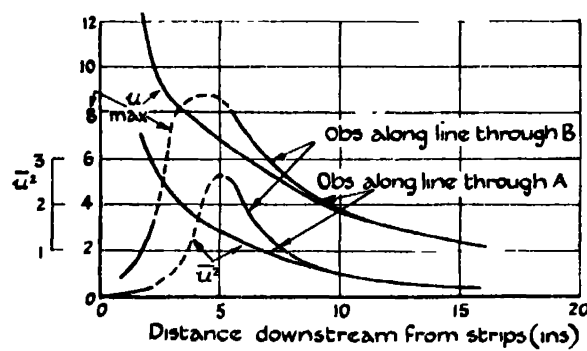


FIG. 26.

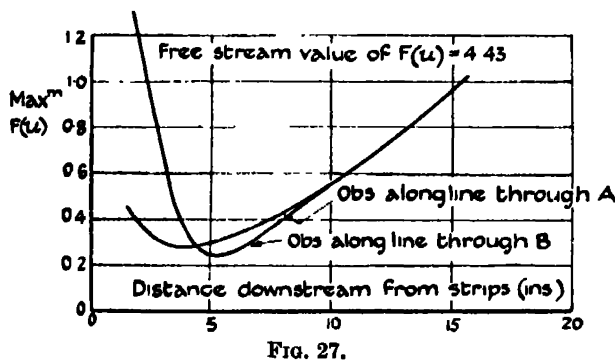


FIG. 27.

FIGS. 25-27.—Velocity distribution.

Table II.—Summary of Results.

Position of wire downstream (inches).	Local mean wind speed (ft./sec.).	Range of $u$ (ft./sec.).	$u_{max}$	$\frac{\int u dx}{\delta}$ (secs.)	$h$ .	$\frac{1}{2h^2}$	Max. $P(u)$ .	$u^2$ .	$\frac{u_{max}}{\sqrt{u^2}}$
Along line through A.									
2	4.1	+8.7 -2.85	11.55	64.3	—	—	0.387	3.15	6.5
3	3.84	+5.85 -2.6	8.45	69	—	—	0.294	2.172	5.74
6	5.19	+3.5 -2.7	6.2	61	0.584	1.465	0.331	1.232	5.6
9	5.65	+1.94 -2.13	4.07	64.5	0.868	0.663	0.49	0.565	5.42
15	5.29	+1.15 -1.16	2.31	62.4	1.68	0.176	0.95	0.162	5.74
Along line through B.									
	15.3	+1.18 -1.21	2.39	63	2.142	0.109	1.21	0.164	5.9
3	13.4	+2 -5.6	7.6	58.6	1.185	0.355	0.67	0.418	5.73
6	7.2	+4.15 -3.17	7.32	63	0.478	2.19	0.271	2.016	5.16
9	5.76	+2.46 -1.70	4.22	62.5	0.832	0.72	0.48	0.614	5.4
15	5.29	+1.15 -1.15	2.3	58.5	1.68	0.176	0.95	0.162	5.74
Free Stream }	5.29	{ +0.265 -0.27 }	0.535	73	7.85	0.0081	4.43	0.0077	6.1

consists of irregular disturbances having a maximum amplitude of about 16 to 20% of the corresponding eddying disturbances, and a frequency of about 50 per second against 55 per second for the eddy frequency.

Lastly, reference should be made to the results obtained from a statistical analysis of some of the records, based on the number of times the velocity variations attain certain selected values during the time occupied by the experiment. For obvious reasons, the analysis was confined to records in which the wave form of the disturbances could be definitely identified; therefore, except for one record extending over a period of 5 seconds, all the measurements were made with the faster records like those of figs. 9-12, Plate 6. The frequency distribution curves plotted from the results were found to be irregular, except for the longer record, relating to the flow at 15 inches back from the grid. This proved to be identical with the curves shown previously to be associated with completely diffused turbulent motion and obeying the normal error law.

In conclusion, the authors desire to acknowledge their indebtedness to Professor G. I. Taylor for his advice and co-operation, and also for the loan of a suitable oscillograph for carrying out the experiments.

### *Summary.*

Records of turbulence of an air stream made by an Einthoven oscillograph recording the variations in current through a hot wire are often difficult to analyse statistically. Recently a method has been developed for determining both the frequency distribution function  $F(u)$  and the root-mean-square velocity  $\bar{u}$ , by amplifying the out-of-balance potential variations generated across the Wheatstone bridge which contains the hot wire, and recording them by allowing the light reflected from the mirror of a balanced-armature oscillograph to fluctuate to and fro across a stationary photographic plate for a period of 4 or 5 minutes. When developed the plate reveals a dark image of variable density on a light background, the distribution of which, measured photometrically, provides a means of estimating the total equivalent time of exposure at all points along the record, and hence data from which velocity frequency curves can be plotted and values of  $\bar{u}$  computed.

The method was applied to the analysis of free stream turbulence in a 1-foot wind tunnel, and to the eddying wake of a grid composed of horizontal and vertical strips, held across the tunnel about 3 feet from the inlet. All experiments were undertaken at a wind speed of 5.3 feet per second; with the grid in position explorations were made along two lines parallel to the tunnel axis, one passing through the centre of a hole in the grid, the other through a neighbouring strip.

Values of the average velocity  $V_0$ ,  $F(u)$ ,  $u_{\max}$ , and  $\bar{u}$  convey some idea of the diffusion of the eddy motion from the wake of a strip into the high velocity jet through the neighbouring hole. The process continues for a distance of about 10 inches downstream when turbulence becomes uniformly distributed across the section. Under such conditions the frequency distribution is found to resemble closely that of the "normal error law," namely,

$$F(u) = \frac{h}{\sqrt{\pi}} e^{-h^2 u^2}.$$

A similar law is also found to hold for the small scale turbulence of the free stream; the values of  $h$  and  $\bar{u}$  are 7.85 and 0.088 respectively, whereas at 10 inches downstream of the grid the corresponding figures are 1.68 and 0.40.

---

*An Analysis of the Fine Structure of the  $\alpha$ -particle Groups from Thorium C and of the Long Range Groups from Thorium C'.*

By W. B. LEWIS, B.A., and B. V. BOWDEN, B.A., Goldsmith Senior Student.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received December 29, 1933.)

During the course of our previous measurements of the fine structure of  $\alpha$ -particle groups\* we obtained preliminary measurements of the fine structure of the  $\alpha$ -particle groups from thorium C. The results were in general agreement with those previously obtained by Rosenblum and Valadares,† who photographed the  $\alpha$ -ray spectrum. Our method of measurement appeared capable of yielding results of accuracy as high as 1 in 100,000 in velocity. Accordingly a series of measurements has now been carried out to examine the thorium C groups with much greater precision, and these experiments are described in the present paper. Our results provide information, which we believe to be of considerable accuracy, concerning the energy levels of the thorium C nucleus. When interpreted by Gamow's‡ theory these measurements are in excellent agreement with certain  $\gamma$ -ray energies deduced from entirely independent  $\beta$ -ray measurements.§

By accelerating the particles in an electrostatic field we have in addition been able to measure the energy difference between the two most prominent groups of  $\alpha$ -particles from thorium C, by a direct electrostatic measurement. This independent measurement is probably the most accurate determination of this energy difference, and is in excellent agreement with the energy of a strong  $\gamma$ -ray deduced from the  $\beta$ -ray spectrum measurements by Ellis.

The long range groups of particles from thorium C' were re-examined in detail. This examination confirmed our earlier conclusion that only two groups exist. The new measurements are in substantial agreement with the  $\gamma$ -ray energies corresponding to the  $\beta$ -rays recently discovered by Sze and by Ellis.||

\* Rutherford, Wynn-Williams, Lewis, and Bowden, 'Proc. Roy. Soc.,' A, vol. 139, p. 617 (1933), quoted later as I.

† 'C. R. Acad. Sci. Paris,' vol. 194, p. 967 (1932).

‡ 'Nature,' vol. 126, p. 397 (1930).

§ Ellis, 'Proc. Roy. Soc.,' A, vol. 138, p. 318 (1932).

|| 'Proc. Roy. Soc.,' A, vol. 143, p. 350 (1934).



*Experiments.*

The experimental arrangement of the annular magnet, slit system, ionization chamber, amplifying and counting apparatus was substantially the same as previously described.\* The control and measurement of the magnetic field were greatly facilitated by an automatic control developed by Wynn-Williams, p. 250. This mechanism continuously corrected for slow changes in the magnet field. These changes might be due to slight changes in the temperature of the magnet or to changes in the energizing circuit. The elimination of these fluctuations allowed the operator to concentrate his attention on other experimental details. As described in our previous paper, the fluxmeter has been used as a null instrument, and has been kept as nearly as possible at its natural zero. Any residual thermo-electric e.m.f. in the fluxmeter circuit results in a slow change in the flux through the circuit. Precautions were taken to make this as small as possible, and the resulting rate of drift does not amount to more than the equivalent of a change of about 4 gauss in the magnet gap in 2 hours. Since the flux in the magnet gap is of the order of 10,000 gauss this change is small, but it must be allowed for in all these measurements. This can readily be done by returning to a previously observed group, which may be located to within 1 part in 100,000 of the magnetic field.

*Sources.*—The radioactive sources used in these experiments have been thorium active deposit collected on a polished non-magnetic stainless steel plate, 8 mm.  $\times$  3 mm.  $\times$  0.4 mm. thick. The edges of the plates were carefully cleaned with emery paper after exposure. The sources were mounted in the magnet so that their projected width in the direction of emission of the  $\alpha$ -particles observed was only about 0.7 mm. The strength of the sources required in these measurements varied from 0.2 to 1.5 mg. radium equivalent, according to the groups to be observed, and the time of preparation of the sources varied from 2 to 18 hours. Most of the sources used showed no trace of tarnish, and the polish of the surface remained perfect. The sharpness of the peaks obtained confirms that these sources were remarkably clean—see figs. 1 and 2—and shows a considerable improvement on those previously obtained, I, *loc. cit.*, using sources deposited on platinum wires.

*Thorium C.*—It may be recalled that Rosenblum discovered that the  $\alpha$ -particles from thorium C were emitted in five groups, which he named in the order of their intensities  $\alpha_0$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ , and  $\alpha_4$ . Of these the group  $\alpha_1$  is the highest energy group and is about a third of the intensity of the adjacent group  $\alpha_0$ . The remaining groups are of much weaker intensity. For our experiments on

\* I (*loc. cit.*).

these groups the resolving power of the apparatus was made as great as possible by the use of a narrow counter slit 0.43 mm. wide, and also by limiting the spread of the selected  $\alpha$ -particle beam by a defining slit to a maximum of 2 cm. at the mid-point of the trajectory.

A special method was employed to measure the energy difference between the two main groups thorium C  $\alpha_0$  and  $\alpha_1$ . In our first paper we stated that it would be very difficult to measure this energy difference by our usual method with greater accuracy than 1 part in 200, corresponding to about 200 electron volts, which would involve measuring the field to within a tenth of a gauss in a field of about 10,000 gauss. We have, in fact, been able to make this measurement with rather greater accuracy than this, but it seemed desirable to measure this difference by an entirely different method, namely, by electrostatic acceleration of the particles as illustrated in fig. 7 of our first paper.\* If the magnet field is maintained constant at a value midway between those appropriate to the two groups, it is possible to bring the two groups successively on to the counting chamber by accelerating or retarding the  $\alpha$ -particles. The energy difference between the two groups is about 40,000 electron volts, and as the charge on the  $\alpha$ -particle is  $2e$ , a potential of 10,000 volts is required. The advantage of this method is that it is possible to pass instantly from one group to the other, so that there is no error due to the drift of the magnetic field, and all magnetic measurements are completely eliminated. The energy difference is measured directly in volts, and as our result is expressed in electron volts, it is not necessary to assume any physical constants in this determination. It is therefore an absolute measurement, and may be contrasted with our other measurements in which we have had to adopt a value for the velocity and energy of a standard group of  $\alpha$ -particles (see footnote to Table III).

It will be realized that it is difficult to obtain a vacuum of the highest order in the interior of our magnet. Nevertheless, a sufficiently high vacuum was obtained to withstand the potential applied, but occasional discharges were liable to occur when the source was negatively charged. The discharge current was, however, limited by a series resistance of 30 megohms, so that the surface of the source was not appreciably affected by the discharges. It was at first found difficult to apply a potential of 10,000 volts to the source owing to the breakdown of the insulation of the source holder. This difficulty was overcome by making a holder of a single piece of thick walled pyrex capillary tubing into which the high voltage lead was sealed.

\* I (*loc. cit.*).

The potential was generated by six 2000-volt Evershed generators connected in series, as these were available in the laboratory. The two driving motors were run from accumulators, and the output voltage was constant to about 50 volts. The small fluctuations were eliminated by means of a screened grid valve stabilizer,\* which reduced them to less than 5 volts. Since it was necessary to reverse the potential, and there were unavoidable small leaks in the generators to earth, it was essential that the current stabilizer should always be connected to the high potential terminal of the generators. The potential was measured in terms of the current through a 50-megohm resistance, so it was necessary to avoid any leak of the stabilized current of the order of 0.2 microamps. to earth, which implies an insulation resistance of the order of  $5 \times 10^{10}$  ohms. This was achieved by the use of metallic shields connected to the high potential end of the generators, which reduced the possibility of direct leak to earth of the stabilized current to a minimum.

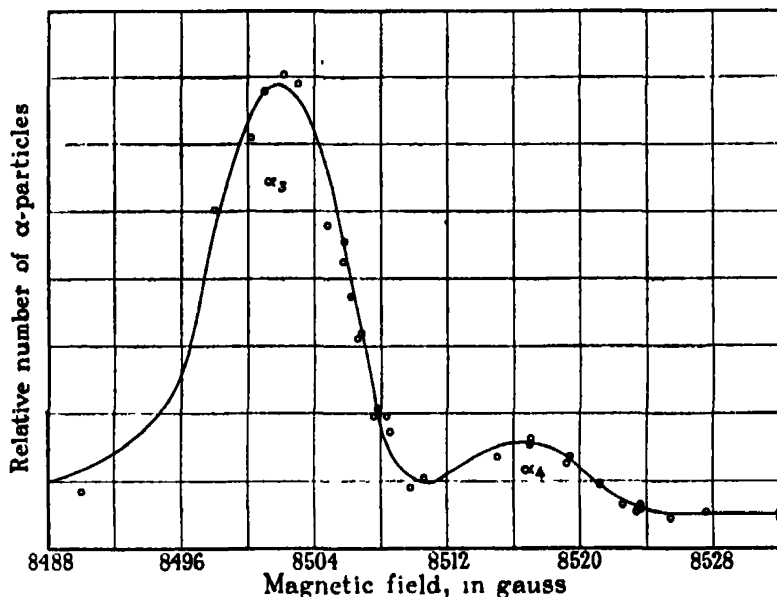


FIG. 1.—Thorium C,  $\alpha_3$ , and  $\alpha_4$  showing points obtained in one experiment.

The 50 megohm resistance consisted of five selected 10 megohm "Siemens" grid leak resistances immersed in oil. These resistances consist of a carbon deposit in the form of a spiral on a small porcelain tube. Their behaviour has been described in detail by Hartman and Dossman.† These resistances

\* Street and Johnson, 'J. Franklin Inst.,' vol. 214, p. 155 (1932).

† 'Z. Tech. Physik,' vol. 9, p. 434 (1928). See also Bowdler, 'J. Inst. Elect. Eng. Lond.,' vol. 73, p. 66 (1933).

are remarkably constant, and do not vary by more than one or two parts in 5000 under load. They were frequently calibrated by comparison with a standard megohm. The current passing through them was measured by a Tinsley potentiometer, using a Weston standard cell. We are confident that the potential was always known to within 1 part in 1000.

During an experiment there might be a slow drift of the magnet field, and it was therefore necessary to make alternate counts on the two groups in rapid succession. The auxiliary apparatus for this purpose is described by Wynn-Williams.\* This apparatus automatically made alternate counts of 9 seconds each on the two groups over a period of several minutes. The drift during a period of 10 seconds would be less than one part in two million in the total field and can therefore be neglected.

In order to analyse the results, the ratio of the number of particles in the two groups must be known with great accuracy. We made special experiments to determine this ratio, during which we counted about a quarter of a million particles. We conclude that the ratio is  $2.57 : 1 \pm \frac{1}{2}\%$ . The method of reducing the observations was essentially as follows. The shape of the peak of the group thorium  $C\alpha_0$  was determined in a preliminary experiment for each source, by magnetic and electrostatic exploration as illustrated in fig. 7 of our first paper. From these measurements the shape of the peak thorium  $C\alpha_1$  was calculated.

Each experimental count provided a point on  $\alpha_0$  and a point on  $\alpha_1$ , separated by a known energy difference. Each point on  $\alpha_0$  was fitted to the experimentally determined curve, and in this way a series of points on  $\alpha_1$  was obtained, to which the calculated curve could be fitted. The deviations of the points from this curve were consistent with the fluctuations to be expected when counting a limited number of particles. A sufficient number of particles was counted to make any such statistical errors small compared with other possible systematic errors. In the absence of discharges the current through the 30 megohm safety resistance was found to be negligible, so that the full measured potential was applied to the source. Further, by calculation of the electrostatic field, verified by experiment, it was found that the  $\alpha$ -particles were accelerated within such a small distance from the source that no error arises from considering the acceleration to take place at the source itself. It was necessary to take precautions to ensure that the full potential was applied to the source for the whole counting period. The change over of the applied voltage was sufficiently rapid to allow a half second interval before the count

\* 'Proc. Phys. Soc. London,' vol. 46, p. 303 (1934).

was resumed. During this half second the soaking of the charge into the insulation was practically completed, and any error from this cause was small. Any uncorrected error from this cause would tend to make the measured energy difference greater than the true value.

The final result for the difference between the two groups is 39.31 electron kilo-volts, corresponding to a difference in energy of disintegration of 40.07 e.k.v. after correcting for the energy of the recoiling nuclei. If any of the above mentioned causes of error have been underestimated the result will tend to be higher than the true value. The result is therefore probably not low, and should be accurate to 1 part in 400.

This energy difference was also measured as accurately as possible by the ordinary magnetic method. The result for the disintegration energy difference is 39.90 e.k.v. and the error is probably not greater than 1 part in 300. We do not consider the difference between the two determinations to be significant, and adopt the weighted mean value 40.0<sub>0</sub> e.k.v.

The measurement of the weak groups  $\alpha_2$ ,  $\alpha_3$ ,  $\alpha_4$  which differ in velocity by a maximum of 4% from the main groups necessitates an accurate knowledge of the behaviour of the magnetic field.

It must be recalled that the magnetic field is measured at one point, and it is necessary to find the relation of this measured field to the effective mean field over the semicircle traversed by the  $\alpha$ -particles. It was anticipated, I, *loc. cit.*, p. 627, that the circumferential variations of the field would depend to some extent on the previous magnetic history of the iron, and not only was this found to occur, but it was also found that these variations were quite comparable with the very small total inhomogeneity. It is therefore necessary to know at what point on the hysteresis figure of the magnet it is being operated.

These variations were investigated by connecting a movable search coil, of approximately equivalent area turns to the coils used for measuring the field, in opposition to these coils in the fluxmeter circuit. The movable coil was mounted on a framework so that it could be moved round the semicircle traversed by the  $\alpha$ -particles. By this arrangement small changes in the field as a whole had no effect on the fluxmeter, and the exploration of the inhomogeneity could be made with great accuracy. Differences between the measured and the effective mean fields could be determined to better than 0.1 gauss, or 1 part in 100,000 in the total field. Greater accuracy than this was not required, as it was found that after changes of the magnet field of the order of 4%, differences in the inhomogeneity of about 0.2 gauss were observed even when the magnetic history of the magnet as a whole was the same in both

experiments. This fact imposes a limit to the accuracy obtainable by the present method, and in the present experiments the observational errors have been reduced so that these indeterminable variations of the inhomogeneity are the greatest remaining source of uncertainty in the results. Nevertheless, the accuracy of measurement of these energy differences is comparable with the accuracy of the corresponding  $\beta$ -ray measurements. This implies an accuracy of 1 part in 30,000 in these determinations of the relative velocities of the  $\alpha$ -particle groups. It would be extremely difficult to make  $\beta$ -ray measurements more accurately than it is possible to measure  $\alpha$ -particle energy differ-

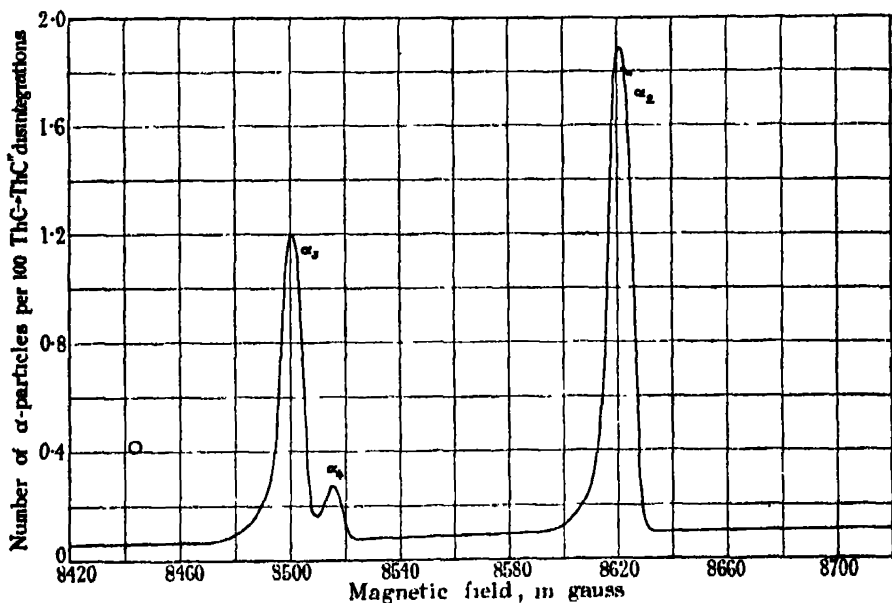


FIG. 2.

ences, owing to the fact that  $\beta$ -ray lines are always superimposed on a very large continuous background.

Our results, which are included in Table III, are illustrated in fig. 2, which shows the groups  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ . The curve can be resolved into three peaks of similar shape. The experimental points have not been included, as on this very small scale, each point would appear to lie exactly on the curve. The small continuous background is due to retarded particles from the two large groups  $\alpha_2$  and  $\alpha_1$ . It will be observed that the height of this background is considerably less than that of the peak  $\alpha_2$ , which illustrates the great improvement on our preliminary measurements (*cf.* fig. 8 of our first paper). There is a

satisfactory separation between the two groups  $\alpha_3$  and  $\alpha_4$ , which differ in velocity by only 1.5 parts in 1000.

We made special experiments to look for the group  $\alpha_5$  tentatively suggested by Rosenblum,\* but we obtained no evidence for the existence of this group, and conclude that any such group must be less than 1 in 7000 of the group  $\alpha_0$ ; in other words, the number of particles emitted in this group is less than 1 in 10,000  $\alpha$ -particle disintegrations of thorium C.

*Long Range Thorium C'.*—Owing to the fact that the strongest sources available were only of about 2 mg. radium equivalent, the number of particles in these groups was very small, and we were obliged to use the same experimental conditions as those used in our analysis of the long range particles from radium C'.† This unfortunately involved the loss of the highest resolving power of the apparatus.

Our results for the weaker group are shown in fig. 3, the group appears to be single and there is no evidence of an accompanying group as suggested by Rosenblum.\* We can be quite certain that if any such group exists it must be less than one-twentieth of the observed group. Owing to the much greater number of particles counted, these measurements are much more reliable than those described in our first paper. Moreover, a correction has now been applied for the hysteresis variation of the inhomogeneity mentioned above. Measurements of the group of longest range have been repeated, and are in good agreement with our previous measurements.

### Discussion.

The theory of the fine structure of the  $\alpha$ -rays from thorium C was first put forward by Gamow.‡ Summarizing the current views it is simply assumed that all the parent thorium C atoms and all the final thorium C'' atoms have the same energy and that the total energy emitted in the disintegration is the same whatever the intermediate processes may be. If, therefore, an  $\alpha$ -particle is emitted with lower energy than this total energy, the resulting nucleus is left in an excited state, and the balance of energy is emitted in the form of one or more quanta of  $\gamma$ -radiation. We denote these energy states of the thorium C'' nucleus corresponding to  $\alpha$ -particle emission in the groups  $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4$  by  $E_0, E_1, E_2, E_3, E_4$  respectively. Whatever the mechanism of the process may be, the theory predicts that the energies of the  $\gamma$ -rays corresponding to

\* Rosenblum and Valadarez, *loc. cit.*

† Rutherford, Lewis, and Bowden, 'Proc. Roy. Soc.' A, vol. 142, p. 347 (1933).

‡ 'Nature,' *loc. cit.*

transitions from the excited states  $E_0, E_2, E_3, E_4$  should be equal to the energy differences between the corresponding  $\alpha$ -particle disintegrations and the normal disintegration, which may be identified with the group  $\alpha_1$ . In some cases two or more quanta of  $\gamma$ -radiation may be emitted in succession. This involves the existence of transition states which may or may not be the same as the states excited in the nucleus by the disintegration of thorium C. Table I includes the energy differences deduced from our results, and the  $\gamma$ -rays deduced from independent measurements by Ellis of the corresponding  $\beta$ -ray lines. It will be noted that the agreement between corresponding values is within the probable limits of experimental error in each instance.

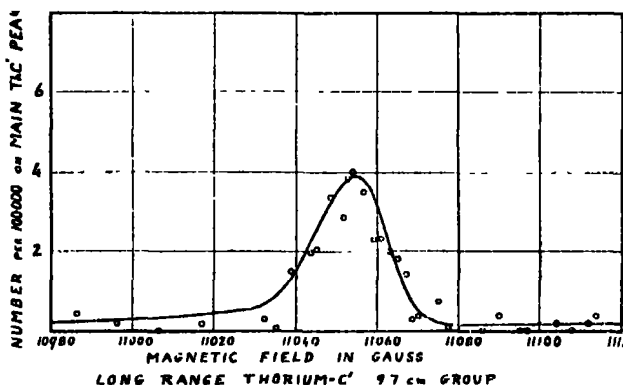


FIG. 3.—Long range thorium C' 9.7 cm. group.

Table I.—Energies in electron kilo-volts.

	Our measurements.	$\gamma$ -rays.
$E_0-E_1$	40.0 <sub>0</sub>	39.95
$E_2-E_1$	327.8	326.7
$E_3-E_1$	472.4	470.9
$E_4-E_1$	491.8	—
$E_2-E_0$	287.8	286.9
$E_3-E_0$	432.4	431.7
$E_4-E_0$	451.8	451.1
$E_1-E_2$	144.6	—
$E_1-E_4$	19.4	—

It is implicit in the theory that for each  $\alpha$ -particle which is emitted in one of the lower energy groups, a  $\gamma$ -ray quantum will be emitted in a transition from the corresponding excited energy state of the thorium C'' nucleus. The counting method of observing the  $\alpha$ -particles provides very definite information of their intensity. Since the  $\gamma$ -rays indicate transitions only into the levels



$E_0$  and  $E_1$ , the number of quanta produced by transitions from  $E_2$ ,  $E_3$ ,  $E_4$  will be equal to the number of  $\alpha$ -particles in the corresponding groups, and since the transitions into  $E_0$  are few compared with its frequency of excitation by the  $\alpha$ -particle disintegration, the number of quanta from the transition  $E_0$ - $E_1$  may be expected to be equal to the number of  $\alpha$ -particles in the group  $\alpha_0$  within the limits of accuracy of  $\gamma$ -ray intensity measurements. This question has been discussed in detail by Ellis and Mott,\* and until further data on the probable values of the internal conversion coefficients of  $\gamma$ -rays are available, there is not much that can be added to their discussion. It seems likely, however, that the  $\alpha$ - and  $\beta$ -ray intensities may rather be used to determine the actual internal conversion coefficients by interpretation according to the theory outlined above. Since it appears that the internal conversion coefficient varies very markedly with the associated change of nuclear azimuthal quantum number,† it may then be possible to deduce probable quantum numbers of the excited energy levels from the theory of internal conversion, which is, at the moment, still under revision.‡

It must not be forgotten that  $\gamma$ -ray data alone are not adequate to determine uniquely the energy levels of a radioactive nucleus, since the  $\gamma$ -rays correspond only to the energy differences between levels, and in general these may be combined in a variety of ways to form self-consistent level systems. Measurements of  $\alpha$ -ray fine structure, on the other hand, provide information of the energy levels which is definite and unambiguous, though transition energy levels may exist which may be deduced from the  $\gamma$ -rays, and which are not indicated by the  $\alpha$ -particle measurements. We found it necessary to postulate such levels in our consideration of the levels of the excited radium C' nucleus.§

Although the groups  $\alpha_2$ ,  $\alpha_3$ ,  $\alpha_4$  contain relatively few  $\alpha$ -particles, the frequency of excitation of an  $\alpha$ -particle into the level from which it escapes may be very much greater, for it is well known that the rate of  $\alpha$ -particle escape varies rapidly with the energy of the emergent  $\alpha$ -particle. The rate of escape is therefore considerably less for these groups than for the higher energy groups

\* 'Proc. Roy. Soc.,' A, vol. 139, p. 369 (1933).

† Taylor and Mott, 'Proc. Roy. Soc.,' A, vol. 138, p. 665 (1932); vol. 142, p. 215 (1933).

‡ It should be noted that the intensity of the group  $\alpha_4$  is less than was assumed by Ellis and Mott in their discussion of the levels of the thorium C-C' nucleus. Our measurements confirm the agreement of the energy difference  $E_4$ - $E_0$  with the associated  $\gamma$ -ray, so it seems probable either that the intensities of the very weak  $\beta$ -ray lines have been over-estimated, or that the internal conversion coefficient of the  $\gamma$ -ray is larger than was assumed.

§ Rutherford, Lewis, and Bowden, *loc. cit.*

$\alpha_1, \alpha_0$ . We may suppose that the rate of  $\alpha$ -particle escape is a function both of its emergent energy and the energy and quantum number of the resulting excited nucleus. A large number of variables may therefore be involved, but it seems reasonable to assume that the contribution due to the dependence of the rate of  $\alpha$ -particle escape on the disintegration energy may be estimated from Gamow and Houtermans'\* formula assuming a constant "radius" for the nucleus. When the numbers of particles in the groups are divided by the appropriate disintegration constants calculated in this way it appears that the resulting quotients are all of the same order of magnitude as is shown in Table II. It is probable that the excitation probabilities for all the levels in the thorium C nucleus from which  $\alpha$ -particles escape, are therefore also of the same order of magnitude. This follows if the other variables do not have a very marked effect on the rate of disintegration.

Table II.

Group.	$\frac{V_{\alpha_n}}{V_{\alpha_1}}$	$(\text{Disintegration constants})^{-1}$	$\frac{N_{\alpha_n}}{N_{\alpha_1}}$	Relative excitation probability.
$\alpha_1$	1.0000	1.00	1.00	1.00
$\alpha_0$	0.9968	1.48	2.57	3.80
$\alpha_3$	0.9733	31	0.0663	2.05
$\alpha_4$	0.9614	153	0.0058	0.9
$\alpha_5$	0.9597	197	0.0404	8.0

From their analysis of the available data, Ellis and Mott (*loc. cit.*) have suggested that the levels  $E_0, E_2, E_4$  have quantum number 2, while the level  $E_3$  probably has quantum number 3. This is not in disagreement with the result shown in Table II that the level  $E_3$  would appear to have an abnormally large excitation probability if it were assumed to have the same quantum number as the other three levels.†

From our revised measurements of the two long range groups from thorium C' it is to be expected that  $\gamma$ -rays of energies  $7.26$  and  $17.97 \times 10^5$  electron volts should exist corresponding to transitions of the excited thorium C' nucleus to the ground state before its disintegration. Ellis has observed weak  $\beta$ -rays corresponding to  $\gamma$ -rays of energies  $7.26$  and  $18.02 \times 10^5$  electron volts, which are in good agreement with the expected energies.

\* 'Z. Physik,' vol. 52, p. 496 (1929).

† [Note added in proof, February 10, 1934.—Gamow and Rosenblum ('C. R. Acad. Sci. Paris,' vol. 197, p. 1620 (1933)), have considered this problem from a slightly different aspect, and they suggest as the result of a calculation essentially similar to that outlined above, that the emission of group  $\alpha_3$  is associated with the lowest quantum number, and that  $\alpha_1$  and  $\alpha_4$  both correspond to some high quantum number.]

For convenience of reference the velocity and energy data of all accurately measured  $\alpha$ -particle groups are collected in Table III.

We wish to express our thanks to Lord Rutherford for his advice and direction throughout the whole course of this work, which continues the analysis of  $\alpha$ -rays which we did in collaboration with him. We also wish to thank Dr. Wynn-Williams for his collaboration in much of the experimental work, and Dr. Ellis for much helpful discussion of its theoretical aspects, and for the use of the radioactive material, part of which was given to him by Professor Schlundt of the University of Missouri, and part of which was purchased by a grant from the Royal Society. We are indebted to the Department of Scientific and Industrial Research, and to the Goldsmith's Company for their grants to enable us to carry out this work.

### *Summary.*

The fine structure of the  $\alpha$ -particle groups from thorium C has been examined in detail by counting methods employing magnetic focussing as previously described. These measurements, which are believed to be more accurate than previous determinations, confirm very exactly the previously established agreement with  $\gamma$ -ray energies required by Gamow's theory of the fine structure.

The energy difference between the two main groups has been measured directly in electron volts by alternately accelerating and retarding the  $\alpha$ -particles by an applied potential of about 10,000 volts, which was accurately measured. In this way the two groups were alternately made to have the same velocity, and were examined without alteration of the magnetic field or movement of the source or counter. The result is in close accord with the energy difference measured by the usual method of altering the magnetic field, and also with the independent  $\beta$ -ray measurements of the corresponding  $\gamma$ -ray energy.

In addition the energies and relative intensities of all the fine structure  $\alpha$ -ray groups have been measured. The two long range groups from thorium C' have been re-examined with greater accuracy than before. The excess energies of these groups above the normal group are in good agreement with the  $\gamma$ -ray energies calculated from the very weak  $\beta$ -ray lines measured by Ellis.

No very reliable discussion of the level system of the thorium C-C'' nucleus can yet be given, owing to possible changes in the theory of internal conversion of  $\gamma$ -rays. Our results, however, appear to confirm the conclusions reached by Ellis and Mott.

A table which summarizes all the accurate data of  $\alpha$ -ray emitted by radioactive substances is given at the end of the paper.

Table III.

$\alpha$ -particle group.	Mean range cm. of air at 15° C. 760 mm.	$V/V_0$				Velocity cm./sec. $\times 10^{-4}$	Energy of $\alpha$ -particle e. volts $\times 10^{-4}$	Energy of disin- tegration e. volts $\times 10^{-4}$	Difference of energy of dis- tegration from main group e. volts $\times 10^{-4}$	Relative number of $\alpha$ -particles.	Basis for velocity and energy data.
		Annular magnet.	Briggs.	Rosen- blum.	Deduced from range.						
Radium $\alpha_0$	(3.26)	—	—	0.789 <sub>4</sub>	—	1.517	4.793	4.879	0	—	Ros.
" $\alpha_1$	(3.08)	—	—	0.774 <sub>3</sub>	—	1.488	4.612	4.695	1.84	—	Ros.
Radon "	4.014	0.8448 <sub>0</sub>	0.8455 <sub>18</sub>	0.844 <sub>3</sub>	0.8452	1.62512	5.4879 <sub>2</sub>	5.5896 <sub>7</sub>	—	—	B.
Radium A	4.620	0.8837 <sub>0</sub>	0.8940 <sub>17</sub>	0.8836 <sub>3</sub>	0.8836	1.69910	6.0002 <sub>4</sub>	6.1123 <sub>9</sub>	—	—	B.
" C $\alpha_0$	(4.039)	0.8470	—	—	—	1.6279	5.506 <sub>3</sub>	5.611 <sub>7</sub>	0	94	M.
" C $\alpha_1$	(3.969)	0.8423	—	—	—	1.6189	5.445 <sub>3</sub>	5.549 <sub>9</sub>	0.622	113	M.
" C'	6.870	1.00000	1.00000	1.0000	1.0000	1.922 <sub>90</sub>	7.683 <sub>90</sub>	7.829 <sub>44</sub>	0	10 <sup>4</sup>	1. (footnote)
Long range groups from radium C'	7.755	1.0380	—	—	1.037 <sub>9</sub>	1.9850	8.260	8.437	6.08	0.43	M.
	9.00	1.0785	—	—	—	2.0729	8.941	9.112	12.83	(0.45)	M.
	—	1.0861 <sub>4</sub>	—	—	1.086 <sub>1</sub>	2.0876	9.068 <sub>9</sub>	9.241 <sub>6</sub>	14.12	22	M.
	—	1.1008	—	—	—	2.1157	9.315	9.493	16.63	0.38	M.
	—	1.1111	—	—	—	2.1356	9.492	9.673	18.44	1.35	M.
	—	1.1209	—	—	—	2.1543	9.660	9.844	20.15	0.35	M.
	—	1.1279	—	—	—	2.1678	9.781	9.968	21.38	1.06	M.
	—	1.1351	—	—	—	2.1817	9.908	10.097	22.68	0.36	M.
	—	1.1447	—	—	—	2.2001	10.077	10.269	24.39	1.67	M.
	—	1.1488	—	—	—	2.2079	10.149	10.342	25.13	0.38	M.
	—	1.1589	—	—	—	2.2274	10.329	10.526	26.97	1.12	M.
	11.47	1.1689	—	—	1.168 <sub>3</sub>	2.2466	10.509	10.709	28.80	0.23	M.
Polonium	3.805	0.8310 <sub>9</sub>	—	0.8308 <sub>4</sub>	0.8310	1.5971 <sub>9</sub>	5.300 <sub>3</sub>	5.403 <sub>3</sub>	—	—	M.
Radio actinium $\alpha_0$	—	—	—	0.987 <sub>9</sub>	—	1.706 <sub>1</sub>	6.051	6.159	0	80	Ros.
" $\alpha_1$	—	—	—	0.885 <sub>7</sub>	—	1.702 <sub>1</sub>	6.019	6.127	0.32	15	Ros.
" $\alpha_2$	—	—	—	0.883 <sub>5</sub>	—	1.697 <sub>9</sub>	5.990	6.097	0.62	100	Ros.
" $\alpha_3$	—	—	—	0.881 <sub>5</sub>	—	1.694 <sub>9</sub>	5.968	6.075	0.84	15	Ros.

Table III—(continued).

$\alpha$ -particle group.	Mean range cm. of air at 15° C. 760 mm.	$V/V_e$				Velocity cm./sec. $\times 10^{-4}$ .	Energy of $\alpha$ -particle e. volts $\times 10^{-4}$ .	Energy of disin- tegration e. volts $\times 10^{-4}$ .	Difference of energy of disin- tegration from main group e. volts $\times 10^{-4}$ .	Relative number of $\alpha$ -particles.	Basis for velocity and energy data.
		Annular magnet.	Briggs.	Rosen- blum.	Deduced from range.						
Radio actinium $\alpha_1$	—	—	—	0.878 <sub>0</sub>	—	1.688 <sub>5</sub>	5.924	6.030	1.29	5	Ros.
" $\alpha_2$	—	—	—	0.874 <sub>5</sub>	—	1.680 <sub>0</sub>	5.870	5.975	1.84	10	Ros.
" $\alpha_3$	—	—	—	0.870 <sub>5</sub>	—	1.673 <sub>0</sub>	5.817	5.921	2.38	5	Ros.
" $\alpha_4$	—	—	—	0.866 <sub>5</sub>	—	1.665 <sub>0</sub>	5.766	5.869	2.90	80	Ros.
" $\alpha_5$	—	—	—	0.862 <sub>5</sub>	—	1.662 <sub>0</sub>	5.744	5.847	3.12	15	Ros.
" $\alpha_6$	—	—	—	0.863 <sub>5</sub>	—	1.658 <sub>9</sub>	5.719	5.822	3.37	60	Ros.
" $\alpha_7$	—	—	—	0.859 <sub>0</sub>	—	1.652 <sub>4</sub>	5.674	5.776	3.83	10	Ros.
Actinium X $\alpha_8$	—	—	—	0.863 <sub>5</sub>	—	1.658 <sub>0</sub>	5.719	5.823	0	6	Ros.
" $\alpha_9$	—	—	—	0.864 <sub>5</sub>	—	1.642 <sub>0</sub>	5.607	5.709	1.14	4	Ros.
" $\alpha_{10}$	—	—	—	0.849 <sub>0</sub>	—	1.631 <sub>4</sub>	5.533	5.634	1.89	1	Ros.
Actinium $\alpha_0$	5.655	—	—	0.942 <sub>7</sub>	0.9415	1.811 <sub>7</sub>	6.826	6.953	0	10	Ros.
" $\alpha_1$	(5.30 <sub>1</sub> )	—	—	0.924 <sub>7</sub>	—	1.776 <sub>5</sub>	6.561	6.683	2.70	1	Ros.
" $\alpha_2$	(5.14 <sub>7</sub> )	—	—	0.915 <sub>4</sub>	—	1.769 <sub>5</sub>	6.436	6.556	3.37	1	Ros.
" $\alpha_1$ & $\alpha_2$ (mean)	5.203	—	—	0.919 <sub>0</sub>	0.9175	—	—	—	—	2	—
Actinium A	6.420	—	—	0.980 <sub>2</sub>	0.9794	1.882 <sub>4</sub>	7.368	7.508	—	—	R.
Actinium C $\alpha_0$	5.392	—	—	0.9286	0.9277	1.7332	6.611	6.739	0	100	R.
" $\alpha_1$	4.947	—	—	0.9040	0.9030	1.7356	6.262	6.383	3.56	19	R.
Actinium C'	6.518	—	—	0.984 <sub>5</sub>	0.9839	1.8911	7.437	7.581	—	0.32	R.
Radio thorium $\alpha_0$	—	—	—	0.840 <sub>5</sub>	—	1.615 <sub>0</sub>	5.420	5.517	0	5	Ros.
" $\alpha_1$	—	—	—	0.833 <sub>5</sub>	—	1.602 <sub>0</sub>	5.335	5.431	0.86	1	Ros.
Thorium X	—	—	0.86042	0.860 <sub>2</sub>	—	1.653 <sub>7</sub>	5.6825	5.7858	—	—	B.

Thorium	4.967	0.90464	0.9046	0.9041	1.7387	6.2832	6.399 <sub>s</sub>	—	B.
Thorium A	5.601	0.93985	0.939 <sub>s</sub>	0.9388	1.8084	6.7759	6.903 <sub>s</sub>	—	B.
Thorium C $\alpha_1$	—	—	—	—	—	—	—	—	M.
" $\alpha_2$	—	0.89014	0.8906	—	1.7108 <sub>s</sub>	6.083 <sub>s</sub>	6.200 <sub>s</sub>	27.2	M.
" $\alpha_3$	—	0.88727	0.8876	—	1.7053 <sub>s</sub>	6.044 <sub>s</sub>	6.160 <sub>s</sub>	69.8	M.
" $\alpha_4$	—	0.86635	0.8664	—	1.6651 <sub>s</sub>	5.762 <sub>s</sub>	5.872 <sub>s</sub>	1.80	M.
" $\alpha_5$	—	0.85584	0.8556	—	1.6445 <sub>s</sub>	5.620 <sub>s</sub>	5.728 <sub>s</sub>	0.16	M.
" $\alpha_6$	—	0.85420	0.8541	—	1.6417 <sub>s</sub>	5.601 <sub>s</sub>	5.708 <sub>s</sub>	1.10	M.
" $\alpha_6$ & $\alpha_7$ (mean)	4.693	0.8889 <sub>s</sub>	0.8881	0.8881	—	—	—	—	—
Thorium C'	8.533	1.0687 <sub>s</sub>	1.0690	1.0687	2.0540 <sub>s</sub>	8.778 <sub>s</sub>	8.947 <sub>s</sub>	10 <sup>6</sup>	M.
" L.R.	9.687	1.1110 <sub>s</sub>	1.109 <sub>s</sub>	1.111 <sub>s</sub>	2.1354 <sub>s</sub>	9.491 <sub>s</sub>	9.673 <sub>s</sub>	34	M.
" L.R.	11.543	1.1707 <sub>s</sub>	1.170 <sub>s</sub>	1.170 <sub>s</sub>	2.2501 <sub>s</sub>	10.541 <sub>s</sub>	10.744 <sub>s</sub>	190	M.

1. From the measurements of Briggs ('Proc. Roy. Soc., A, vol. 118, p. 549 (1928)) and of Rosenblum and Dupouy ('C. R. Acad. Sci. Paris,' vol. 194, p. 1919 (1932)) we previously assumed the energy of disintegration of radium C' to be  $7.8293 \times 10^6$  electron volts. This value has been confirmed by Rosenblum's latest absolute measurement of the velocity of these particles (Rosenblum and Dupouy, 'J. Phys. Rad.,' vol. 4, p. 262 (1933)) and it was also adopted by Briggs ('Proc. Roy. Soc., A, vol. 139, p. 638 (1933); vol. 143, p. 604 (1934)). This value depends both on the value of  $E/M$  and the accurate determination of  $H_p$  for the  $\alpha$ -particle and is probably correct to about 1 part in 1000, and all values of energies and of energy differences are subject to a systematic uncertainty of this magnitude. A greater number of significant figures than this assumption justifies is given since the measurements of relative velocities are considerably more accurate.

2. In the last column the measurement used as a basis for the calculation of the energies is indicated. "M" = the velocity measured with the annular magnet. "B" = velocity measured by Briggs. "Ros." = Rosenblum. "R" = mean range.

3. The values of  $V/V_0$ , the ratio of the velocity of the  $\alpha$ -particle group to the velocity of the main  $\alpha$ -particle group from radium C' obtained from all four sources are given.

4. All the velocity ratios and energies have been corrected for relativity effects.

5. The ranges given in brackets have not been measured, but have been deduced from the velocity given.

6. References—See footnote 1, and Rosenblum and Chamue, 'C. R. Acad. Sci. Paris,' vol. 196, p. 1598 (1933). Curie and Rosenblum, *ibid.*, vol. 196, p. 1598 (1933). Rosenblum, *ibid.*, vol. 195, p. 317 (1932). Rosenblum and Valadares, *ibid.*, vol. 194, p. 967 (1932). Rosenblum, *ibid.*, vol. 193, p. 848 (1931). Rutherford, Lewis, and Bowden, 'Proc. Roy. Soc., A, vol. 142, p. 347 (1933).

7. In the actinium series the velocities have been calculated from the energies given by Rosenblum.

8. Our recent re-exploration of the field of our magnet has enabled us to make very slight corrections to some of our previously published results.

## *An Automatic Magnetic Field Stabilizer of High Sensitivity.*

By C. E. WYNN-WILLIAMS, Ph.D.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received January 11, 1934.)

The present paper describes a method whereby the field of an electromagnet can be kept constant automatically to at least one part in fifty thousand (*i.e.*, to 0.2 gauss in 10,000 gauss) for long periods. The method has the great advantage of not merely ensuring that the current through the magnet (or, say, the voltage across the magnet windings) is kept constant, but that the *field itself*, as measured by a fluxmeter,\* is kept within one-fifth of a gauss (or less) of some desired value. Thus, it is not necessary to know whether a fluctuation of the field is caused by unsteadiness of the exciting battery voltage, variations of the magnet coil resistance with temperature, etc.; in all cases, the stabilizer automatically readjusts the exciting current until the field is brought back to its original value. Further, as explained at the end of the paper, the stabilizer can be made to bring about an increase or decrease in the field of a definite number of gauss, and to continue to maintain the field at this new value. Also with very slight modifications, which are so obvious as to require no further explanation, the stabilizer could be made to maintain a current, or a potential difference, constant just as easily as a magnetic field.

The method has been applied with success to the annular magnetic field used in the Cavendish Laboratory.† The magnet is used for the analysis of  $\alpha$ -particle groups. Under the influence of suitable fields of the order of 10,000 gauss,  $\alpha$ -particles of various velocities are made to travel in a semicircle of 40 cm. radius, and are focussed on to the slit of a counting chamber. Groups of various velocities are focussed in turn by varying the field, their velocities being determined from the values of the fields. It is extremely important, therefore, to know the field strength with accuracy and to maintain it constant, at each of these values, during the experiment. In some experiments already described‡ this was done manually. An observer watched the spot of light

\* It should be pointed out, however, that this sensitive automatic control is merely applied to a method of field measurement previously described. The accuracy of control therefore suffers from the same limitation of a possible slow continuous drift of the magnetic field. [Rutherford, Wynn-Williams, Lewis and Bowden, 'Proc. Roy. Soc.,' A, vol. 139, p. 617 (1933).]

† Cockcroft, 'J. Sci. Instr.,' vol. 10, p. 71 (1933).

‡ *Loc. cit.* and Rutherford, Lewis and Bowden, 'Proc. Roy. Soc.,' A, vol. 142, p. 347 (1933).

reflected from the mirror of a fluxmeter, the coil of which was connected to a stationary search coil situated in the magnetic field. A change of one-fifth of a gauss in the magnetic field produced a deflection of the fluxmeter spot of about 1 mm., and by suitable adjustment of the field current by means of a mercury rheostat it could be arranged that the spot was kept within half a millimetre of its zero position. This, however, required one observer to spend his whole time in stabilizing the field during an experiment which might last several hours.

In more recent experiments,\* the automatic stabilizer described here has been used to replace the observer, and has proved perfectly satisfactory. Throughout the experiments—lasting several hours—magnetic fields of the order of 10,000 gauss were kept stabilized to within 0.2 gauss.

The general method adopted consists of replacing the observer by a double cathode photo-electric cell, each cathode of which is connected through the medium of an extremely simple valve and thyatron circuit (operated entirely from the 200-volt alternating current mains, without any rectification) to a system of relays. Movement of the fluxmeter spot actuates the relays, causing an electric motor rapidly to re-adjust the magnet current in a similar manner to that used by the observer. Experiments showed, however, that in order to secure the appropriate form of correction for various kinds of field disturbance likely to be experienced, and to guard against accidental failure of the current supplies, it was essential to employ a rather more intricate system of relay connections than at first seemed necessary. Several possible photo-cell and thyatron circuits of a simple kind were tried, with some measure of success, before the present form was finally adopted. These methods are mentioned at the end of the paper.

*Thyatron Circuits.*—Fig. 1, c, shows the photo-cell and thyatron circuit used, and fig. 2 the auxiliary relay circuits. Figs. 1, a, and 1, b, are simplified circuits added for the purpose of more easily explaining the action of that of fig. 1, c.

Referring to fig. 1, a, T is a thyatron, whose anode circuit is supplied with alternating current from the secondary,  $S_1$ , of the transformer. Normally, an arc is prevented from striking in the thyatron (and no anode current passes) during a positive half cycle because the grid potential is kept more negative than the critical value by the grid being supplied, through the resistance R, with an out-of-phase potential from the opposed secondary winding,  $S_2$ , of the same transformer. When the cathode of the photo-cell is illuminated, a

\* Lewis and Bowden, p. 235.



current passes during the positive half cycles of anode potential. This causes the thyatron grid potential to become less negative, as it sets up a potential difference in  $R$  which opposes the bias supplied from  $S_2$ . If this current is sufficiently great, the thyatron grid bias becomes more positive than the critical bias, and an arc strikes during each positive half cycle. Thus, a mean anode current commences to flow suddenly at some degree of illumination of the cell. Such a simple circuit, however, proved to be insufficiently sensitive for the purpose.

Fig. 1, *b*, shows how the system is improved by the addition of one stage of valve amplification. The valve is merely substituted for the photo-cell, its

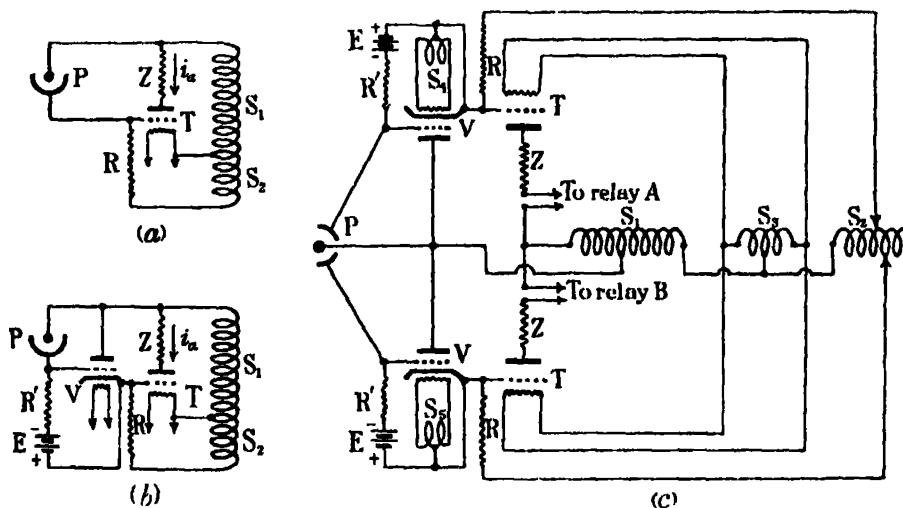


FIG. 1.—In 1, *c*.  $P$  = photo-cell.  $T$  = Mazda B.T.1 thyatron.  $V$  = Mullard 904V valve.  $S_1$  = 200 volts, with tap at 100 volts.  $S_2$  = 20 to 30 volts.  $S_3$  = 2 volts.  $S_4$  and  $S_5$  = 4 volts.  $E$  =  $4\frac{1}{2}$  volts.  $R$  = 100,000 ohms.  $R'$  = 10 megohms.  $Z$  = 450 ohms.

anode and cathode being connected to the points to which those of the photo-cell were previously connected, while the anode and cathode of the photo-cell are connected to the anode and grid of the valve. An increase in the cell illumination causes a greater anode current to flow in the valve during the positive half cycle, which in turn opposes the bias of the thyatron. The thyatron anode current commences to flow suddenly at a degree of illumination of the photo-cell, which is considerably lower and more sharply defined than it was with the simpler circuit.

By using two photo-cell circuits similar to the above (or one double cathode photo-cell), it can be arranged that a movement of the fluxmeter spot from

zero, in either direction, will cause one circuit or other to function (thus indicating that the field is too high or too low) and to apply a suitable correction through the medium of the auxiliary apparatus which is described later. Fig. 1, *c*, shows the complete double circuit employed for this purpose, which is that illustrated in fig. 1, *b*, arranged in a practicable form. All the potential and current supplies are obtained from one transformer, with the exception of the grid bias for the two valves, which is obtained from two small dry batteries. It is necessary to employ screened leads between the photo-cell and the valves, and, for stability, one valve (with its associated thyatron) is screened from the other. It is necessary, also, to be able to vary the grid bias of the two thyatrons independently of each other, in rather fine steps; this is effected by suitable tapings on winding  $S_2$ . There is a tendency for the striking or extinction of the arc in one thyatron to affect the other thyatron circuit; this can be avoided by careful adjustment of each of the valve and thyatron grid bias potentials; a further improvement might possibly be effected by using separate transformers for the two circuits so as to reduce interaction between them.

The photo-electric cell is housed in an adjustable earthed metal case under the fluxmeter scale; adjustment of the stabilizer zero is effected by a screw which moves the cell in a direction parallel to the scale. The image of a rectangular aperture, about a centimetre wide, is focussed by the fluxmeter mirror (which is about 4 mm. square) on to the photo-cell, so as to cover a portion of each of the cathodes. The slit is illuminated by a condenser and an overrun 12 volt 48 watt lamp.

In an arrangement such as the above, two methods of working are possible. In the first, and more obvious method, it can be arranged that when the field is normal, both thyatrons are "out"; that when it is too high, one thyatron, A, is alight, and when too low, the other thyatron, B, is alight. In the alternative method, it can be arranged that when the field is normal, *both* thyatrons are alight; that A goes out when the field is slightly too high, and that B goes out when it is slightly too low. The latter method is much the superior, since, in any case, a large change of field will move the spot right off the cell and cause *both* thyatrons to go out. In the latter method this is different from the "normal field" state, but not in the former. The various indications which are possible with this latter arrangement are:—

- (1) Field normal . . . . Both thyatrons A and B alight.
- (2) Field slightly high . . Thyatron A extinguishes, B remains alight.
- (3) Field very high . . . Thyatron B extinguishes after A extinguishes.

- (4) Field slightly low      Thyatron B extinguishes, A remains alight.  
 (5) Field very low . . . . Thyatron A extinguishes after B extinguishes.

Thus, there are *five* different indications available, as compared with only three with the first arrangement, and the most distinctive state (the normal) is a "positive" one.

*Auxiliary Apparatus.*—In order to use these five indications for effecting suitable corrections, auxiliary apparatus must be used. Indications (3) and (5), both of which have the same final state (both thyratrons A and B out), require that the apparatus should be capable of distinguishing between the two states, and that it should be able to continue to apply the appropriate form of correction after the second thyatron has gone out; further, that simultaneous accidental extinction of both thyratrons caused by a lamp failure, or failure of the alternating current supply, should not cause the stabilizer to operate incorrectly and so disturb the field.

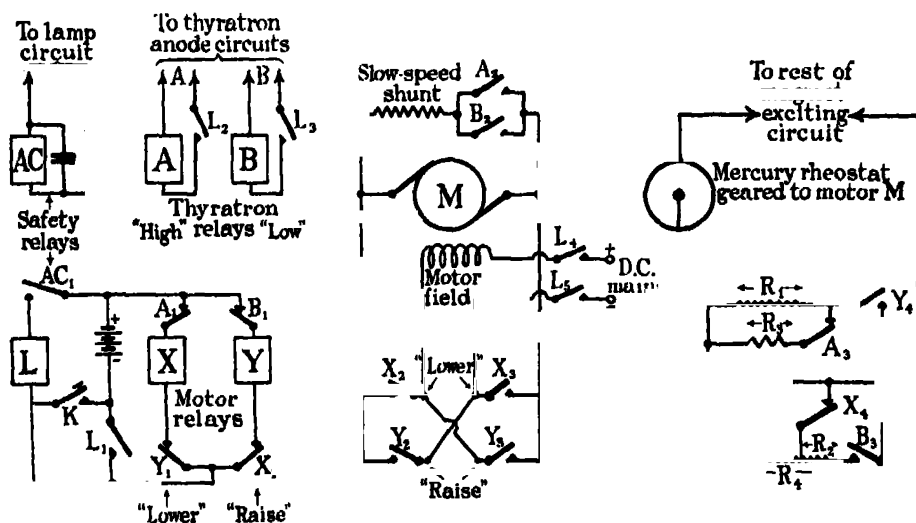


FIG. 2.

The translation of the five indications into suitable current corrections is done as follows. Referring to fig. 2, M is a series motor, coupled through suitable gearing to the mercury rheostat which adjusts the electro-magnet current. Current is supplied to the motor armature through two pairs of relay contacts,  $X_2$  and  $X_3$ , and  $Y_2$  and  $Y_3$ , which have their connections reversed, and which are operated by two relays X and Y. Operation of X causes the motor to lower the field and Y to raise it. Relays X and Y are operated respectively by contacts  $A_1$  and  $B_1$  of two relays A and B through the windings of which

pass portions of the anode currents of thyratrons A and B. In the normal state, when the thyratrons are alight, relays A and B are operated, and the circuits of X and Y are broken by contacts  $A_1$  and  $B_1$  being open. When either thyatron goes out, the associated relay releases and causes X or Y, as the case may be, to operate and drive the motor forwards or backwards to adjust the field. It is impossible for *both* relays X and Y to operate together, even if both thyratrons are out and contacts  $A_1$  and  $B_1$  both closed; for whichever relay, X or Y, operates first, opens the circuit of the other at contact  $X_1$  or  $Y_1$ , and so prevents it from operating. Bearing this in mind, the five states appear as relay combinations, as follows:—

	Thyatron relays.		Motor relays.	
	Relay A.	Relay B.	Relay X.	Relay Y.
(1) Field normal . . . . .	Operated	Operated	Free	Free
(2) Field slightly high . . . .	Free	Operated	Operated	Free
(3) Field very high . . . .	Free	Free	Operated	Free
(4) Field slightly low . . . .	Operated	Free	Free	Operated
(5) Field very low . . . . .	Free	Free	Free	Operated

From this it will be seen that the motor continues to run in the correct directions at stages (3) and (5), even though the fluxmeter spot may be right off the scale. In such circumstances, it is desirable for the correction to be made rather more quickly than is necessitated at stages (2) and (4). This is easily arranged by connecting a pair of contacts  $A_2$  and  $B_2$  in parallel with each other, to apply a shunt resistance across the motor armature. This causes it to run slowly when *either* of the relays A or B is in the operated state (at stages (2) or (4)); when *both* are free, however (at stages (3) or (5)), the shunt is broken, and the motor speed is considerably increased.

*Checking of "Hunting."*—Owing to the large self-inductance of the magnet, there is a tendency for the field to "overshoot" the normal value after a correction has been applied; this causes the mechanism to "hunt." This, however, can be checked as follows. The main exciting current of the magnet is passed through a low resistance R. Contacts  $X_4$  and  $Y_4$  operated by relays X and Y arrange that immediately either relay operates a sudden small correcting change (a decrease for X and an increase for Y) is made in the magnet current by switching the shunt resistances  $R_1$  or  $R_2$  in or out of circuit. This sudden change takes place before the motor moves, and has the effect of "advancing" the motor correction, so that a smaller movement of the rheostat takes place. When the spot of the fluxmeter reaches the zero position, relay

X or Y releases, the motor stops and a small sudden change of magnet current now occurs in the reverse direction. This serves to check the tendency of the field to overshoot the normal value, and so "hunting" is reduced. Similarly, two other resistances,  $R_3$  and  $R_4$ , are arranged in connection with contacts  $A_3$  and  $B_3$ , so that a rather greater sudden change is made at stages (3) and (5), when the motor speed is greater. The values of the resistances and the motor speeds are adjusted by trial and error until the best working conditions are obtained.

The above arrangement reduces hunting very considerably. It is still further reduced by allowing a slight "overlap" of thyratrons A and B; that is, by allowing a margin on either side of the zero position, within which both thyratrons may remain alight. In practice this margin can be reduced to about half a millimetre on either side of the zero with ease, corresponding to a "tolerance" of plus or minus one-tenth of a gauss in the value of the field. Decreasing the "overlap" increases the sensitivity of the stabilizer at the expense of increased "hunting," and *vice versa*. The overlap width is easily controlled by varying the width of the illuminated slit, or more conveniently, by altering the lamp current.

*Safety Circuit.*—Relays L and AC guard the apparatus against failure of current supplies, or burning out of the lamp. The stabilizer is put into operation by closing key K. Relay L operates provided contact  $AC_1$  is closed, and "locks" at contact  $L_1$ , maintaining this state after the key is opened. The remaining L contacts close and complete the stabilizer connections. Should the current supply to the relays fail momentarily, false re-operating of X and Y is prevented by relay L immediately unlocking, and putting the stabilizer out of action until key K is again pressed. Failure of the lamp, or of the A.C. supply to the thyratrons is guarded against by relay AC, through which passes a rectified portion of the alternating current passing through the lamp. When AC releases, the circuit of L is broken at  $AC_1$ , and L puts the stabilizer out of action as before. An alarm is operated under these conditions, so that manual control of the field can be carried out immediately, before it has "drifted" too far for the experiments to be continued.

*Automatic Adjustment of Field.*—Included in series with the search coil and the fluxmeter is a second coil, in the centre of which a permanent magnet can be rotated. This coil, which has been termed a "Compensator" in an earlier paper, is calibrated, so that by rotation of the magnet additional line-turns may be included in the fluxmeter circuit. When it is desired to make a definite change in the magnetic field, *e.g.*, a reduction of 2 gauss, the magnet

of the compensator is turned so as to add the required number of line turns to the fluxmeter circuit. This causes a deflection of the fluxmeter, which in turn causes the stabilizer to commence so to alter the field of the magnet as to return the spot to zero ; this process will be completed when the field has been so reduced that the same number of line turns as were added by the compensator are subtracted at the magnet search coil by the change of field. The compensator scale is conveniently calibrated directly in terms of the magnet scale. Such changes are made slowly, precautions being taken to avoid large deflections of the fluxmeter for reasons given in previous papers.

*Other Possible Methods.*—A simpler and more obvious method of stabilizing was tried by arranging to illuminate the cathode of a single photo-cell, connected with a thyatron according to the well-known "phase control" circuit. Movement of the spot caused an increase or decrease of the mean anode current of the thyatron, which was passed into the magnet windings to assist or oppose the main field current. Such a method works very well, but is not as sensitive as the present arrangement, and, further, it suffers from the great disadvantage of not possessing a definite "zero." Its "zero" is an arbitrary one, defined by the light intensity, thyatron and cell characteristics, supply voltages, etc., and varies during an experiment, thus disturbing the field. The zero of the present arrangement is independent of all these factors, and is "geometrically" defined.

A double circuit phase control arrangement, with and without valve amplification, was also tried, but even this was found to be less satisfactory in operation than the circuit described in this paper, at any rate, for the purpose to which it has been applied.

I wish to thank Mr. Lewis and Mr. Bowden for their help in trying out the various arrangements described in this paper, and for valuable suggestions. I wish also to record my thanks to the British Thomson Houston Company, Rugby, for gifts of suitable thyatrons and photo-electric cells needed for the work.

---

## *X-Ray Examination of Certain Copper-Zinc Alloys at Elevated Temperatures.*

By Professor E. A. OWEN and LLEWELYN PICKUP, M.Sc.(Lond.), Ph.D.(Wales),  
University College of North Wales, Bangor.

(Communicated by Sir William Bragg, O.M., F.R.S.—Received January 9, 1934.)

This investigation is a continuation of the work on copper-zinc alloys reported in previous papers.\* Our previous measurements were made on quenched alloys, and it was assumed that the conditions prevailing at any temperature at the moment of quenching were retained in the quenched samples. In the present work an attempt is made to investigate the validity of this assumption by examining the alloys at the actual temperatures of annealing. Attention is directed mainly to the  $\beta$ -phase in the pure region and in the mixed regions on either side of the pure phase. It was hoped that such measurements might also throw more light on the nature of the  $\beta$ -transformation.

### *Apparatus and Method of Experiment.*

The precision camera was the same in principle as that previously used and described, with modifications in design for high temperature work. It was made entirely of invar except that three silica rods connected the portion of the drum carrying the film to that carrying the sample. By this device the one part was well insulated thermally from the other. In order to take photographs *in vacuo*, the camera was fitted into a brass box with a removable lid and water-cooled sides; inside the box, the heater, consisting of "Kanthal" resistance wire embedded in alundum cement, was mounted. To hold the sample, which was in the form of fine filings on thin foil, against the camera frame, a thin sheet of copper foil was used. This was anchored with insulating porcelain beads and wire springs to the camera frame carrying the film. To minimize the heat passing from the sample to the camera frame, a thin sheet of mica of definite thickness was used around the slot over which the sample was placed. The sample was heated by bringing the heater, shaped to the contour of the camera frame, into close contact with a similarly shaped copper sheet about 1.5 mm. thick, into which the "hot" junction of one thermo-couple was silver-soldered; this in turn pressed against the foil on which the sample was mounted. Leads for the thermo-couple pyrometers—one to

\* 'Proc. Roy. Soc.,' A, vol. 137, p. 397 (1932); vol. 140, pp. 179, 191 (1933).

measure the approximate temperature of the sample photographed and the other to measure the camera temperature near the film—passed through insulated plugs in one side of the box. On the opposite side, leads to the heater were similarly inserted. An outlet for exhausting the box completed the construction.

The exact dimensions of the camera and the fiducial distances were obtained by means of a standard end gauge and a Hilger travelling microscope. These measurements were checked by means of accurate determinations, under definite conditions of experiment, of the parameter values of certain pure metals which have already been well established. The camera was mounted in the box with its plane horizontal; this somewhat lessened the heating of the film but necessitated mounting the X-ray tube with its axis vertical. The same oil pump was used both to obtain the fore-vacuum in the mercury diffusion pump attached to the X-ray tube, and to evacuate the brass box.

One of the thermo-couples above referred to was used to obtain an approximate estimate of the temperature of the sample under investigation, during the time of exposure. To determine this temperature accurately the known thermal expansion of certain standard substances—silver and copper—was used. The procedure was to cover half the foil carrying the reflecting powder, with the sample to be investigated and the other half with silver or copper. Suitable targets were chosen to give lines from both reflecting materials within the range of the camera. By measuring the parameter of the standard, the temperature of the whole sample was readily derived from the dilatation curve of the standard. By this means any errors in temperature measurement were practically confined to errors in the standard curve.

To obtain satisfactory and easily measurable reflection lines on the films from both substances, it was essential to set the camera relative to the X-ray beam very critically. Only a small area of each sample could be irradiated so it was important to obtain as intense a beam as possible from the tube. Also, owing to the readiness with which zinc volatilizes out of brass, the time of exposure at high temperatures had to be kept as short as possible in order to prevent measurable change in the composition of the reflecting sample during exposure. With this type of camera an exposure of 2 hours had hitherto been required to obtain a satisfactory photograph on a double-coated X-ray film of standard make. The exposure in the present work was reduced to half an hour with the aid of a Levy-West intensifying screen pressed uniformly against the film. Satisfactory photographs from "double" samples were not obtained with shorter exposures than this.



Exposures at elevated temperatures *in vacuo* were made in the following manner. The camera, charged with the film and reflecting samples, was inserted in the brass chamber, which was then evacuated to a pressure of about 0.001 mm. Hg. This pressure was maintained for about 1 hour before the heating was commenced to allow the film, which contracted when the chamber was first evacuated, to reach a steady state. The sample was heated during the next 15 minutes to the required temperature, which was given approximately on the galvanometer scale. During the actual exposure, the galvanometer reading was kept strictly constant. The contents of the box were then allowed to cool for about an hour before removing the camera and developing the film.

### *Experimental Results.*

The thermal expansion of the  $\beta$ -constituent in three lump-annealed alloys of the copper-zinc system was examined. These alloys had compositions in the  $(\alpha + \beta)$ , pure  $\beta$ , and  $(\beta + \gamma)$  regions and were marked for reference, 561Z, 523Z, and 482Z respectively. A batch of filings from each was annealed *in vacuo* at 500° C. for 1 hour.

The first set of thermal curves was obtained when silver was used to measure the temperature. The X-ray tube was furnished with a copper anticathode. Tables I, II, and III summarize the experimental results obtained.

It was important to determine the extent of the volatilization of zinc, and its effect on the parameter measurements of both the  $\beta$ -phase and the silver. Some of the samples were therefore re-photographed at atmospheric temperature after being heated *in vacuo*. The order of the changes found can be seen from Table IV.

Table I.—Alloy 561Z.  $\beta$  in the  $(\alpha + \beta)$  region.

Parameter of silver A.	Temperature from curve ° C.	$\beta$ -phase parameter A.	Parameter of silver A.	Temperature from curve ° C.	$\beta$ -phase parameter A.
4.077 <sub>1</sub>	20	2.944 <sub>1</sub>	4.111 <sub>1</sub>	424	2.972 <sub>1</sub>
4.085 <sub>1</sub>	117	2.950 <sub>1</sub>	4.112 <sub>1</sub>	432	2.973 <sub>1</sub>
4.089 <sub>1</sub>	172	2.953 <sub>1</sub>	4.112 <sub>2</sub>	437	2.974 <sub>1</sub>
4.095 <sub>1</sub>	245	2.958 <sub>1</sub>	4.114 <sub>1</sub>	455	2.975 <sub>1</sub>
4.099 <sub>1</sub>	285	2.961 <sub>1</sub>	4.115 <sub>1</sub>	460	2.975 <sub>2</sub>
4.101 <sub>1</sub>	308	2.963 <sub>1</sub>	4.117 <sub>1</sub>	482	2.976 <sub>1</sub>
4.103 <sub>1</sub>	337	2.965 <sub>1</sub>	4.118 <sub>1</sub>	494	2.977 <sub>1</sub>
4.107 <sub>1</sub>	378	2.968 <sub>1</sub>	4.120 <sub>1</sub>	525	2.978 <sub>1</sub>
4.108 <sub>1</sub>	384	2.969 <sub>1</sub>	4.122 <sub>1</sub>	542	2.979 <sub>1</sub>
4.108 <sub>2</sub>	392	2.970 <sub>1</sub>			

Table II.—Alloy 523Z.  $\beta$  in the pure region.

Parameter of silver A.	Temperature from curve ° C.	$\beta$ -phase parameter A.	Parameter of silver A.	Temperature from curve ° C.	$\beta$ -phase parameter A.
4.077 <sub>3</sub>	20	2.945 <sub>1</sub>	4.112 <sub>1</sub>	429	2.974 <sub>9</sub>
4.085 <sub>1</sub>	117	2.950 <sub>7</sub>	4.113 <sub>3</sub>	443	2.975 <sub>9</sub>
4.092 <sub>1</sub>	207	2.956 <sub>3</sub>	4.114 <sub>3</sub>	455	2.976 <sub>8</sub>
4.093 <sub>1</sub>	213	2.956 <sub>6</sub>	4.116 <sub>3</sub>	475	2.978 <sub>3</sub>
4.095 <sub>3</sub>	246	2.959 <sub>3</sub>	4.116 <sub>6</sub>	480	2.979 <sub>3</sub>
4.099 <sub>3</sub>	293	2.962 <sub>7</sub>	4.116 <sub>3</sub>	477	2.978 <sub>7</sub>
4.100 <sub>3</sub>	304	2.963 <sub>3</sub>	4.118 <sub>3</sub>	494	2.979 <sub>7</sub>
4.104 <sub>3</sub>	339	2.966 <sub>6</sub>	4.118 <sub>3</sub>	500	2.979 <sub>3</sub>
4.107 <sub>3</sub>	378	2.968 <sub>3</sub>	4.119 <sub>3</sub>	511	2.980 <sub>3</sub>
4.110 <sub>7</sub>	414	2.972 <sub>3</sub>	4.120 <sub>3</sub>	522	2.980 <sub>7</sub>
4.111 <sub>6</sub>	418	2.972 <sub>3</sub>	4.124 <sub>3</sub>	561	2.983 <sub>1</sub>
			4.124 <sub>3</sub>	561	2.984 <sub>1</sub>

Table III.—Alloy 482Z.  $\beta$  in the ( $\beta + \gamma$ ) region.

Parameter of silver A.	Temperature from curve ° C.	$\beta$ -phase parameter A.	Parameter of silver A.	Temperature from curve ° C.	$\beta$ -phase parameter A.
4.077 <sub>3</sub>	20	2.949 <sub>3</sub>	4.115 <sub>3</sub>	470	2.983 <sub>1</sub>
4.088 <sub>3</sub>	151	2.958 <sub>1</sub>	4.116 <sub>3</sub>	474	2.983 <sub>3</sub>
4.095 <sub>3</sub>	241	2.963 <sub>7</sub>	4.115 <sub>3</sub>	468	2.983 <sub>3</sub>
4.101 <sub>3</sub>	312	2.968 <sub>7</sub>	4.117 <sub>3</sub>	490	2.986 <sub>1</sub>
4.108 <sub>3</sub>	387	2.975 <sub>6</sub>	4.119 <sub>3</sub>	501	2.986 <sub>3</sub>
4.110 <sub>3</sub>	409	2.976 <sub>3</sub>	4.119 <sub>3</sub>	504	2.986 <sub>3</sub>
4.111 <sub>3</sub>	424	2.978 <sub>3</sub>	4.112 <sub>3</sub>	516	2.987 <sub>1</sub>
4.114 <sub>3</sub>	457	2.980 <sub>3</sub>	4.120 <sub>3</sub>	522	2.987 <sub>7</sub>
4.114 <sub>3</sub>	456	2.981 <sub>7</sub>	4.122 <sub>1</sub>	540	2.988 <sub>3</sub>
4.116 <sub>3</sub>	474	2.983 <sub>1</sub>	4.121 <sub>3</sub>	530	2.987 <sub>3</sub>
			4.121 <sub>7</sub>	533	2.988 <sub>3</sub>

Table IV.

Region.	Temperature ° C.	Parameter value (atmospheric temperature).			
		$\beta$ -phase.		Ag.	
		Before.	After.	Before.	After.
$\alpha + \beta$	437	2.944 <sub>3</sub>	2.943 <sub>3</sub>	4.077 <sub>3</sub>	4.076 <sub>3</sub>
$\beta$	339 above 400 417 444 above 600	2.945 <sub>1</sub>	2.945 <sub>1</sub> 2.944 <sub>6</sub> 2.943 <sub>8</sub> 2.944 <sub>3</sub> 2.943 <sub>3</sub>	4.077 <sub>3</sub>	4.077 <sub>1</sub> 4.077 <sub>3</sub> 4.077 <sub>6</sub> 4.077 <sub>3</sub> 4.076 <sub>7</sub>
$\beta + \gamma$	387 540	2.940 <sub>1</sub>	2.948 <sub>1</sub> 2.948 <sub>6</sub>	4.077 <sub>3</sub>	4.076 <sub>6</sub> 4.076 <sub>7</sub>

There is a tendency for the parameter value of each alloy to become slightly less after the heating operation. This suggests a reduction in the concentration of zinc in the reflecting layers. The parameter value of the silver also has a very slight tendency to become less. This could only happen by its dissolving some of the zinc escaping from the alloy. It is to be noted that the changes in the parameters shown in Table IV occurred during the whole period of heating; the amount of change during the actual exposure would therefore be smaller. The effect of these changes is to give too low a value for

Table V.—Alloy 561Z.  $\beta$  in  $(\alpha + \beta)$  region.

Parameter of copper A.	Temperature from curve ° C.	$\beta$ -phase parameter A.	Parameter of copper A.	Temperature from curve ° C.	$\beta$ -phase parameter A.
3.623 <sub>0</sub>	255	2.969 <sub>3</sub>	3.632 <sub>3</sub>	368	2.970 <sub>3</sub>
3.626 <sub>0</sub>	300	2.962 <sub>7</sub>	3.636 <sub>3</sub>	443	2.975 <sub>0</sub>
3.628 <sub>4</sub>	345	2.966 <sub>0</sub>	3.638 <sub>3</sub>	468	2.975 <sub>0</sub>
3.629 <sub>0</sub>	354	2.966 <sub>6</sub>	3.640 <sub>0</sub>	492	2.977 <sub>3</sub>
3.631 <sub>1</sub>	375	2.969 <sub>1</sub>			

Table VI.—Alloy 523Z.  $\beta$  in pure region.

Parameter of copper A.	Temperature from curve ° C.	$\beta$ -phase parameter A.	Parameter of copper A.	Temperature from curve ° C.	$\beta$ -phase parameter A.
3.612 <sub>3</sub>	227	2.958 <sub>7</sub>	3.638 <sub>7</sub>	474	2.978 <sub>0</sub>
3.624 <sub>3</sub>	282	2.962 <sub>3</sub>	3.638 <sub>0</sub>	477	2.979 <sub>0</sub>
3.628 <sub>3</sub>	338	2.966 <sub>3</sub>	3.639 <sub>0</sub>	480	2.979 <sub>7</sub>
3.630 <sub>3</sub>	367	2.969 <sub>0</sub>	3.639 <sub>0</sub>	484	2.979 <sub>0</sub>
3.632 <sub>0</sub>	393	2.970 <sub>3</sub>	3.639 <sub>0</sub>	491	2.980 <sub>0</sub>
3.634 <sub>3</sub>	424	2.974 <sub>3</sub>	3.641 <sub>1</sub>	507	2.980 <sub>0</sub>
3.637 <sub>0</sub>	464	2.979 <sub>0</sub>	3.642 <sub>0</sub>	525	2.982 <sub>0</sub>
3.638 <sub>4</sub>	470	2.978 <sub>0</sub>	3.642 <sub>7</sub>	528	2.982 <sub>4</sub>

the parameter of the  $\beta$ -phase, at a temperature estimated to be slightly owing to the small reduction noted in the silver parameter—than that actually existing during the exposure. To some extent these effects counterbalance one another. But, owing to the importance of the points of inflection shown in the curves when the data in Tables I, II, and III are represented graphically, it was decided to determine portions of these curves by using copper instead of silver to measure the temperature. Any solution of the escaping zinc in the copper would increase its parameter, and therefore the estimated temperature from its standard thermal curve would be slightly too high, exactly the

opposite effect to that produced with silver. The data, when pure copper was used to measure the temperature, are summarized in Tables V and VI. Nickel radiation was used when copper was employed as the standard for temperature measurement.

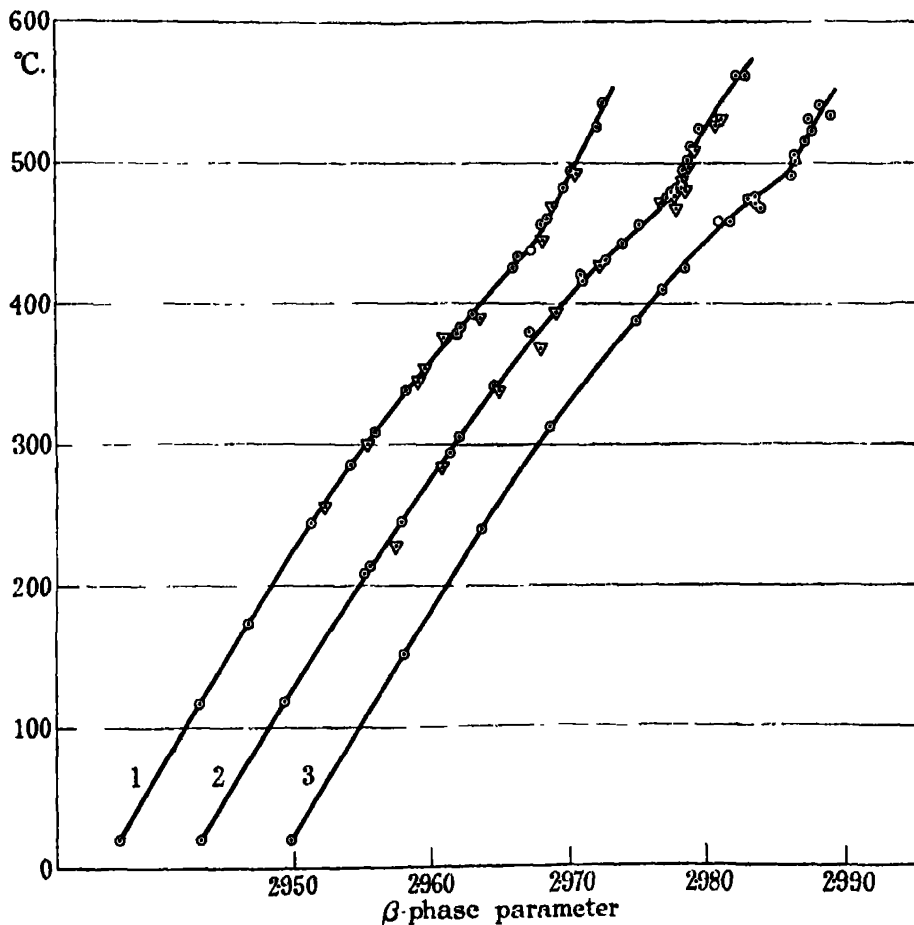


FIG. 1.—Curve 1 shows the expansion in the  $(\alpha + \beta)$  region; Curve 2 refers to the pure  $\beta$ -region, and curve 3 to the  $(\beta + \gamma)$  region. Curves 1 and 2 are displaced to the left relative to curve 3.

From the graphical comparison, fig. 1, of the data in Tables II and VI representing the thermal expansion of the pure  $\beta$ -phase, and those in Tables I and IV for the  $\beta$ -phase in the  $(\alpha + \beta)$  alloy, it is seen that any error resulting from the volatilization of zinc and its solution in silver or copper, is too small to show up definitely in the expansion curves. We are therefore led to conclude that these data represent accurately the thermal expansion of the  $\beta$ -phase.

*Discussion of Results.*

The data from alloy 523Z, which has a composition in the pure  $\beta$ -region represents the expansion of a brass of this composition, while those from alloys 561Z and 482Z represent the expansion of the  $\beta$ -phase constituent in these alloys. We have by X-ray analysis a method of determining the thermal expansion coefficients of each constituent phase independently of the other in an alloy.

Over the portion of each curve up to about 450° C., the coefficients of thermal expansion of the  $\beta$ -phase are found to be practically identical in all three alloys. As shown in our previous work on quenched samples, the phases in the mixed regions are subjected to changes in solid solubility, that is, their

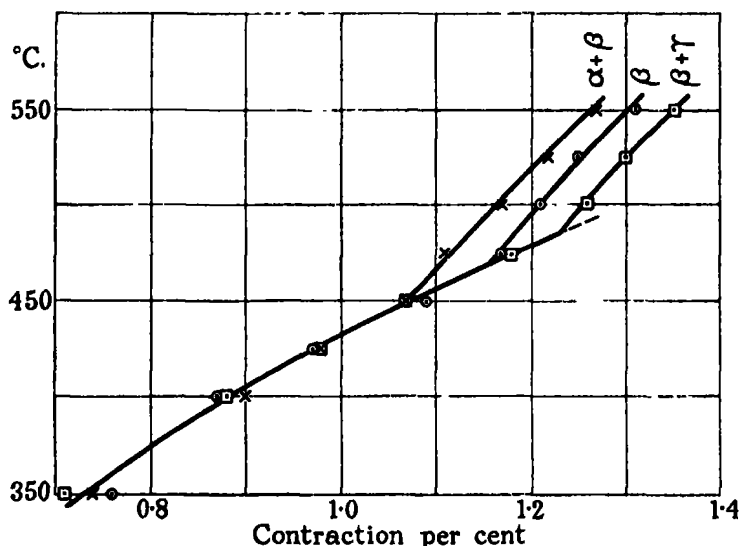


FIG. 2.— $\times$  ( $\alpha + \beta$ ) region ;  $\odot$   $\beta$  region ;  $\square$  ( $\beta + \gamma$ ) region.

saturation composition changes with temperature. It is therefore necessary to eliminate this effect in the  $\beta$ -phase in the ( $\alpha + \beta$ ) and ( $\beta + \gamma$ ) regions, for strict comparison with the  $\beta$ -phase in the pure region. This has been done in Table VII by combining the data on quenched samples already recorded, with the results of the present investigation. In column 4, the difference between the parameter value at the temperature and that after quenching from that temperature is given ; this represents the actual amount of contraction produced during the quenching operation. For better comparison, these differences have been calculated as percentages of the parameter value after quenching from 350° C. in each alloy, and the values are plotted against temperature in fig. 2.

The curves of the three alloys are practically coincident for temperatures below 450° C. ; deviations from the common curve are shown to take place at about 450° C., 470° C., and 485° C. in the curves of the ( $\alpha + \beta$ ),  $\beta$  and ( $\beta + \gamma$ )

Table VII.

Temperature °C.	$\beta$ -parameter value.		Actual contraction on quenching.	Contraction as per cent. of parameter value quenched from 350° C.
	After quenching.	At temperature.		
Alloy 561Z. ( $\alpha + \beta$ ) region.				
550	2.942 <sub>2</sub>	2.979 <sub>2</sub>	0.037 <sub>2</sub>	1.27
525	2.942 <sub>2</sub>	2.978 <sub>2</sub>	0.036 <sub>1</sub>	1.22
500	2.943 <sub>2</sub>	2.977 <sub>2</sub>	0.034 <sub>2</sub>	1.17
475	2.943 <sub>2</sub>	2.976 <sub>2</sub>	0.032 <sub>2</sub>	1.11
450	2.943 <sub>2</sub>	2.974 <sub>2</sub>	0.031 <sub>2</sub>	1.07
425	2.944 <sub>2</sub>	2.972 <sub>2</sub>	0.028 <sub>2</sub>	0.98
400	2.944 <sub>2</sub>	2.970 <sub>2</sub>	0.026 <sub>2</sub>	0.90
350	2.944 <sub>2</sub>	2.966 <sub>2</sub>	0.021 <sub>2</sub>	0.74
Alloy 523Z. ( $\beta$ ) region.				
550	2.944 <sub>2</sub>	2.982 <sub>2</sub>	0.038 <sub>2</sub>	1.31
525	2.944 <sub>2</sub>	2.981 <sub>2</sub>	0.036 <sub>2</sub>	1.25
500	2.944 <sub>2</sub>	2.979 <sub>2</sub>	0.035 <sub>2</sub>	1.21
475	2.944 <sub>2</sub>	2.978 <sub>2</sub>	0.034 <sub>2</sub>	1.17
450	2.944 <sub>2</sub>	2.976 <sub>2</sub>	0.032 <sub>2</sub>	1.09
425	2.944 <sub>2</sub>	2.973 <sub>2</sub>	0.029 <sub>2</sub>	0.97
400	2.944 <sub>2</sub>	2.970 <sub>2</sub>	0	0.87
350	2.944 <sub>2</sub>	2.966 <sub>2</sub>	0.022 <sub>2</sub>	0.76
Alloy 482Z. ( $\beta + \gamma$ ) region.				
550	2.949 <sub>1</sub>	2.989 <sub>2</sub>	0.039	1.35
525	2.949 <sub>1</sub>	2.987 <sub>2</sub>	0.038	1.30
500	2.949 <sub>2</sub>	2.986 <sub>2</sub>	0.037 <sub>1</sub>	1.26
475	2.949 <sub>2</sub>	2.984 <sub>2</sub>	0.034 <sub>2</sub>	1.18
450	2.949 <sub>2</sub>	2.981 <sub>2</sub>	0.031 <sub>2</sub>	1.07
425	2.949 <sub>2</sub>	2.978 <sub>2</sub>	0.028 <sub>2</sub>	0.98
400	2.950 <sub>1</sub>	2.976 <sub>2</sub>	0.025 <sub>2</sub>	0.88
350	2.950 <sub>2</sub>	2.971 <sub>2</sub>	0.021 <sub>2</sub>	0.71

alloys respectively. These temperatures correspond roughly with the transformation temperatures for the different alloys recorded in the equilibrium diagram of the system.

Since the effects on the parameter value due to the change in solubility with temperature have been eliminated, these curves can be taken to represent the

changes in parameter value due to thermal expansion alone. Beyond the temperatures noted above, the alloys have a new thermal coefficient, which appears to have approximately the same value for the  $\beta$ -phase in all three alloys, and which is different from that below about 450° C. Therefore, in addition to the solubility effect, there is, at the temperature of the  $\beta$ -transformation, another effect which reveals itself in the present investigation as an abrupt change in the thermal coefficient. This occurs not only in the adjacent mixed regions but also in the pure  $\beta$ -phase. We conclude then that there is an effect other than that of solubility operating, which may have a more direct bearing on the  $\beta$ -transformation.

As the quenching is considered to prevent any atoms from passing into or out of the lattice during rapid cooling, there seems little doubt that the change in the thermal coefficient now observed must be brought about by some change in the atoms within the lattice. It is evident that all the actual conditions prevailing at the moment of quenching are not retained in the quenched samples. We may interpret the different rate of thermal expansion above the transformation from that below as indicating that the condition within the lattice is different at the two temperatures. Since the lattice structure is the same, either the orientation or the arrangement of the atoms within the lattice must be different. In our previous experiments we looked into the possibility of a change in orientation of the atoms in packing, which was suggested from a consideration of the change in the mean atomic volume which occurs in the different phases along the diagram when a zinc atom replaces a copper atom, but we came to the conclusion that this explanation of the effect was improbable. We are led, therefore, to look for an explanation of the  $\beta$ -transformation in the distribution of the copper and zinc atoms in the lattice. W. L. Bragg suggests that the  $\beta$ -transformation marks the transition from a random distribution of atoms in the lattice to a more ordered distribution. Our results hitherto have not been contrary to this view, but further investigation is necessary before any definite statement can be made as to the exact nature of the transformation.

We wish to express our indebtedness and thanks to the Department of Scientific and Industrial Research for a grant to enable us to carry out the work described in this paper.

#### *Summary.*

A method is briefly described by which accurate values of crystal parameters are obtained from X-ray photographs taken when the material examined is

maintained at high temperature *in vacuo*. The thermal expansions of the  $\beta$ -constituent in the  $(\alpha + \beta)$ ,  $\beta$  and  $(\beta + \gamma)$  regions of the copper-zinc alloy system are found up to 600° C. The X-ray photographs record, in addition to the true thermal expansion, any change in the parameter value produced by a change in solid solubility with temperature. By means of the data in previous work on quenched samples, such changes due to solubility were eliminated, so that the true thermal expansion was obtained. The  $\beta$ -phase in the three regions was then found to expand at the same rate up to the transformation temperature. The rate of expansion is slower above this temperature, but is again the same in each region. These change points noted in the expansion are shown distinctly, and are at about 450°, 470°, and 485° C. in the  $(\alpha + \beta)$ ,  $\beta$  and  $(\beta + \gamma)$  regions respectively; these temperatures correspond to the transformation temperatures recorded by previous workers by other methods. A comparison with our previous results shows that the quenching operation does not retain all the conditions prevailing at the moment of quenching, but with these alloys it is found to be drastic enough to prevent any atoms passing into or out of the lattice. Therefore any explanation of the abrupt change found in the expansion curves is to be sought in the behaviour of the atoms within the lattice. The possibility of a change in orientation of the atoms *in situ* was considered improbable. It is concluded that the  $\beta$ -transformation is more likely to be due to a change in the distribution of the copper and zinc atoms within the lattice.

---



## *The Theory of the Change in Resistance in a Magnetic Field.*

By H. JONES and C. ZENER, Wills Physics Laboratory, University of Bristol.

(Communicated by R. H. Fowler, F.R.S.—Received January 10, 1934.)

### 1. *Introduction.*

The resistance of a metal is in general increased by a magnetic field. For sufficiently small magnetic fields this dependence may, of course, be expressed by the equation

$$\Delta R/R = BH^2.$$

The calculation of the coefficient  $B$  has been the subject of many previous investigations. Sommerfeld\* has shown that if the electrons are regarded as moving in a constant potential, then the theoretical value of  $B$  is 10,000 times smaller than the observed value, and the temperature dependence is in the wrong direction. Peierls† has suggested that the correct magnitude of  $B$  may be obtained if cognizance is taken of the fact that the electrons are moving in a periodic potential, as is usual in the modern theory of metals. Moreover, he has shown that the correct temperature dependence will then be obtained. Blochinzev and Nordheim‡ have recently investigated in detail the change of resistance of divalent metals from this standpoint. The periodicity of the lattice was, however, introduced in quite an idealized manner. In place of an actual metal they considered a simple cubic lattice. The surface of the Fermi distribution was assumed to be composed of sections each of which was a quadratic function of the quantum numbers  $\xi$ ,  $\eta$ ,  $\zeta$ .§ The coefficient  $B$  was obtained in terms of the Fourier coefficient of the potential energy of an electron in the lattice,  $V_{100}$ . Agreement with the experimental values of  $B$  was obtained by taking  $|V_{100}|$  as small as several hundredths of an electron volt, while the correct value must be of the order of 1 volt. Since in their model  $B$  varies inversely as the square of  $|V_{100}|$ , the comparison with experiment is far from satisfactory.

One purpose of this paper is to develop the theory of the change in resistance with no assumption of this type about the form of the Fermi distribution (§ 2).

\* 'Z. Physik,' vol. 47, p. 56 (1928); Frank and Sommerfeld, 'Rev. Mod. Phys.,' vol. 3, p. 1 (1931).

† 'Leipziger Vorträge' (1930); 'Ann. Physik,' vol. 10, p. 97 (1931).

‡ 'Z. Physik,' vol. 84, p. 168 (1933).

§  $\xi$ ,  $\eta$ ,  $\zeta$  are defined in § 2 of this paper.

The usual assumption that the time between collisions with the lattice is a function only of the energy of the electron, but not of the direction of its motion, is, however, introduced. Secondly, we apply this theory to an *actual* metal (§ 3). The theory is found most readily adaptable to lithium. For this metal we obtain excellent agreement with the observations of Kapitza.

## 2. Theory.

The first part of our problem is to find the distribution function  $f$  of the electrons when both a uniform electric field and a uniform magnetic field are present. We shall consider  $f$  a function of a quantum number  $\mathbf{k}$  having components  $\xi$ ,  $\eta$ ,  $\zeta$ . The eigenfunctions of the electrons unperturbed by the magnetic field are specified by  $\mathbf{k}$  by means of the relation

$$\psi_{\mathbf{k}}(x, y, z) = e^{2\pi i (\xi x + \eta y + \zeta z)} U_{\mathbf{k}}(x, y, z),$$

where  $U_{\mathbf{k}}(x, y, z)$  has the same periodicity as the lattice, namely,  $a$ .

If we now assume that the time between collisions of the electrons with the lattice is a function only of the energy of the electrons, the condition that  $f$  remain stationary in time is\*

$$\frac{ea}{h} \left( \mathbf{F} + \frac{1}{c} \mathbf{v}_{\mathbf{k}} \times \mathbf{H} \right) \cdot \text{grad}_{\mathbf{k}} f + \frac{f - f_0}{\tau} = 0. \quad (1)$$

Here  $\mathbf{F}$  and  $\mathbf{H}$  are the electric and magnetic fields, respectively,  $f_0$  is the undisturbed distribution function,  $\mathbf{v}_{\mathbf{k}}$  is the velocity associated with state  $\mathbf{k}$ , and  $\tau$  is a function only of  $E(\mathbf{k})$ , the energy associated with  $\mathbf{k}$ . The mean free path associated with a state  $\mathbf{k}$  is equal to  $\tau/|\mathbf{v}_{\mathbf{k}}|$ . It is to be particularly noted that this equation is valid only for values of  $\mathbf{H}$  such that†

$$(eH\tau)/mc \ll 1.$$

It is thus sufficient to obtain a solution of (1) as a power series in  $\mathbf{H}$ .

The  $z$  axis will be chosen parallel to  $\mathbf{H}$  so that  $\mathbf{H} = H_z$ . In view of the relation

$$\mathbf{v}_{\mathbf{k}} = (a/h) \text{grad}_{\mathbf{k}} E(\mathbf{k}), \quad (2)$$

\* Blochinzev and Nordheim, *loc. cit.*, equation (2.00).

† Jones and Zener, 'Proc., Roy., Soc.,' A, vol. 144, p. 101 (1934). Blochinzev and Nordheim actually obtain a solution of (1) with their model valid for all values of  $\mathbf{H}$ . In view of the above condition upon equation (1), such a solution should only be applied when  $\mathbf{H}$  is sufficiently small. Their solution, moreover, is only valid when the surface of the Fermi distribution is assumed to be a quadratic function of  $\xi$ ,  $\eta$ ,  $\zeta$ . We must therefore find an approximate solution valid when this is not the case.

we shall introduce the operator

$$\Omega = \frac{\partial E}{\partial \eta} \frac{\partial}{\partial \xi} - \frac{\partial E}{\partial \xi} \frac{\partial}{\partial \eta}.$$

This operator has the properties that if  $S(E)$  is an arbitrary function of  $E(\mathbf{k})$  then

$$\Omega S = S \Omega.$$

Further, we introduce the function  $\Phi(\mathbf{k})$  defined by the equation

$$f = f_0 + \Phi \frac{\partial f_0}{\partial E}. \quad (3)$$

The differential equation (1) may now be written as

$$(eH/c) (a/\hbar)^2 \Omega \Phi + \Phi/\tau + (ea/\hbar) \mathbf{F} \cdot \text{grad}_{\mathbf{k}} E = 0. \quad (4)$$

We have, of course, neglected the product of  $\mathbf{F}$  and  $\Phi$ . This is justified, since  $\Phi$  must contain one component of  $\mathbf{F}$  in every term, and it is never necessary to consider powers of  $\mathbf{F}$  higher than the first.

The solution of (4) as a power series in  $H$  is

$$\begin{aligned} \Phi = - (ea\tau/\hbar) \left\{ \mathbf{F} \cdot \text{grad}_{\mathbf{k}} E - \left( \frac{eH\tau}{c} \right) \left( \frac{a}{\hbar} \right)^2 \Omega \mathbf{F} \cdot \text{grad}_{\mathbf{k}} E \right. \\ \left. + \left( \frac{eH\tau}{c} \right)^2 \left( \frac{a}{\hbar} \right)^4 \Omega^2 \mathbf{F} \cdot \text{grad}_{\mathbf{k}} E - \dots \right\}. \quad (5) \end{aligned}$$

Only when the energy is a quadratic function of  $\xi, \eta, \zeta$ , near the Fermi surface, as in the example of Blochinzev and Nordheim (*loc. cit.*), can this infinite series be written in a closed form.

In the so-called transverse effect the current is measured in the direction of the applied electric field, and is perpendicular to the direction of the magnetic field. Let the applied electric field be in the direction of the  $x$  axis; then according to the above condition the  $y$  component of the current is zero. Therefore by (2) and (3), we have

$$\int \frac{\partial f_0}{\partial E} \Phi \frac{\partial E}{\partial \eta} d\tau = 0, \quad (6)$$

where  $d\tau$  is an element of volume in  $\mathbf{k}$  space. The first approximation to the value of integrals of this kind is obtained by making use of the fact that  $-\partial f_0/\partial E$  possesses a sharp maximum at the value of  $E$  equal to that at the surface of the Fermi distribution at the absolute zero of temperature. Since  $E$  is necessarily an even function of the variables  $\xi, \eta, \zeta$ , it follows that only

those terms of the integrand which are even in  $\xi$ ,  $\eta$ ,  $\zeta$ , will contribute to the integral. Equations (6), (5) give therefore\*

$$\frac{F_x}{F_z} = -\frac{eH\tau}{c} \left(\frac{a}{h}\right)^2 \frac{\int \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial \eta} \Omega \frac{\partial E}{\partial \xi} d\tau}{\int \frac{\partial f_0}{\partial E} \left(\frac{\partial E}{\partial \eta}\right)^2 d\tau}. \quad (7)$$

This equation determines the Hall coefficient when  $E$  is known as a function of  $\mathbf{k}$ .

To find the change in resistance due to the application of the magnetic field we must evaluate the  $x$  component of the current. The number of states in an element  $d\tau$  per cubic centimetre of metal is  $(2N/s) d\tau$ , where  $N$  is the number of atoms per cubic centimetre, and  $s$  the number of atoms per unit cell. Clearly  $(N/s) = 1/a^3$ . The factor 2 appears on account of the electron spin which allows two electrons to occupy each state specified by  $\mathbf{k}$ . Therefore, taking  $f_0(0, 0, 0)$  to be unity, we have

$$i_x = \frac{2e}{a^2 h} \int \frac{\partial f_0}{\partial E} \Phi \frac{\partial E}{\partial \xi} d\tau. \quad (8)$$

Using the solution for  $\Phi$  given by (5) and the value of  $F_y$  given by (7), we obtain

$$i_x/F_x = (2e^2/ah^2) \{ \tau I_1 - \tau^3 (a/h)^4 (eH/c)^2 I_2/I_1 \}, \quad (9)$$

neglecting terms in  $F^2$  and in  $H^4$ .

Here  $I_1$  and  $I_2$  stand for the integrals

$$I_1 = - \int \frac{\partial f_0}{\partial E} \left(\frac{\partial E}{\partial \xi}\right)^2 d\tau \quad (10)$$

$$I_2 = \int \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial \eta} \Omega \frac{\partial E}{\partial \xi} d\tau \int \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial \xi} \Omega \frac{\partial E}{\partial \eta} d\tau - \int \frac{\partial f_0}{\partial E} \left(\frac{\partial E}{\partial \eta}\right)^2 d\tau \int \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial \xi} \Omega^2 \frac{\partial E}{\partial \xi} d\tau. \quad (11)$$

In equation (9) in place of  $\tau$  we may introduce the value of the specific resistance  $R$ , for when  $H$  is zero we have, from (9),

$$1/R = (2e^2/ah^2) \tau I_1.$$

From this relation we obtain at once the expression for the change in resistance for small values of the magnetic field

$$-\Delta i_x/i_x = \Delta R/R = (a^3/2ceR)^2 (I_2/I_1^4) H^2. \quad (12)$$

The quantity  $(I_2/I_1^4)$  is a pure number.

\* When  $E$  is a quadratic function of  $\xi$ ,  $\eta$ ,  $\zeta$ , this expression for the Hall coefficient, becomes identical with that given by Peierls, 'Z. Physik,' vol. 53, p. 263 (1929).

The integral  $I_2$  may be considerably simplified. We observe that if  $G$  is an arbitrary function of  $E(\mathbf{k})$ , then a partial integration gives

$$\int \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial \xi} \Omega G d\tau = - \int \frac{\partial f_0}{\partial E} G \Omega \frac{\partial E}{\partial \xi} d\tau. \quad (13)$$

Hence, if we define two new functions  $p$  and  $q$  as follows

$$p = \left( \frac{\partial f_0}{\partial E} \right)^{\frac{1}{2}} \frac{\partial E}{\partial \eta}$$

$$q = \left( \frac{\partial f_0}{\partial E} \right)^{\frac{1}{2}} \Omega \frac{\partial E}{\partial \xi},$$

and use the notation, that if  $\theta$  denotes any function,

$$(\theta) = \int \theta d\tau,$$

we may write  $I_2$  as

$$I_2 = (p^2)(q^2) - (pq)^2. \quad (14)$$

In alkali metals the surface of the Fermi distribution will not diverge very much from the free electron form, so we write

$$E = \omega |k|^2 + E_2, \quad (15)$$

where

$$\omega = h^2/2ma^2.$$

The function  $q$  may then be written

$$q = 2\omega p + q_2, \quad (16)$$

where

$$q_2 = \left( \frac{\partial f_0}{\partial E} \right)^{\frac{1}{2}} \Omega \frac{\partial E_2}{\partial \xi}.$$

Substituting this value into (14) we obtain

$$I_2 = (p^2)(q_2^2) - (pq_2)^2, \quad (17)$$

which shows that  $I_2$  vanishes when  $E_2$  is zero, i.e., when the surfaces of constant energy are spheres.

In the longitudinal effect the current is measured in the same direction as that in which the magnetic field is applied. The calculation of the change in resistance in this case follows closely the steps of the preceding calculation of the transverse effect. We have

$$i_z = \frac{2e}{a^3 h} \int \frac{\partial f_0}{\partial E} \Phi \frac{\partial E}{\partial \zeta} d\tau,$$

which by (5) leads to

$$\Delta R/R = (a^2/2ceR)^2 (I_3/I_1^3) H^2, \quad (18)$$

where

$$I_3 = \int \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial \zeta} \Omega^2 \frac{\partial E}{\partial \zeta} d\tau.$$

By a partial integration, similar to that which yields (13), this integral reduces to

$$I_3 = - \int \frac{\partial f_0}{\partial E} \left( \Omega \frac{\partial E}{\partial \zeta} \right)^2 d\tau.$$

It follows by Schwarz's inequality that  $I_3$ , given by (14), is positive. Hence in both the transverse and in the longitudinal effect  $B$  is positive. It is easy to see the physical reason for this. The effect of a magnetic field (together with the transverse electric field in the transverse effect) is to diminish the fluctuation of  $f - f_0$  from the zero value in a plane normal to  $H$ . Now before  $H$  is applied  $f - f_0$  is given by

$$f - f_0 = - \tau e F \cdot \mathbf{v}_k \frac{\partial f_0}{\partial E},$$

and is hence proportional to the component of  $\mathbf{v}_k$  along  $F$ , that is, along the direction of the total current. Hence any diminution of the fluctuation of  $f - f_0$  must necessarily decrease the current.

### 3. Application.

In the previous section we have found the expression for the coefficient  $B$  in terms of certain integrals. These integrals may be reduced, by a standard method, to integrations over the surface of the Fermi distribution. In applying our formulæ it is thus advantageous that we select a metal whose Fermi surface has a particularly simple form. Blochinzev and Nordheim (*loc. cit.*) concluded from the consideration of a simple cubic lattice that the Fermi surface of a monovalent metal is of a rather complicated nature, the surface touching the six faces of the cube across which the energy is discontinuous. Now all monovalent metals have either a face or body centred structure. In such a structure the lattice perturbation is much more symmetrical than in a simple cubic lattice, and so it is to be expected that the surface of the Fermi distribution is here simpler than in the simple cubic structure. We show below that this surface is, in fact, nearly spherical, at least in the case of lithium.

Let the Fourier expansion of the potential energy of an electron in a simple cubic lattice be

$$V_e = \sum_{\mathbf{n}} V_{\mathbf{n}} e^{2\pi i \mathbf{n} \cdot \mathbf{r}/a},$$

where  $\mathbf{n}$  stands for a set of three integers. A body-centred lattice is constructed by the superposition of two simple cubic lattices, their relative displacement having the components  $\frac{1}{2}a$ ,  $\frac{1}{2}a$ ,  $\frac{1}{2}a$ . The potential energy of an electron in such a lattice is thus

$$V = \sum_{\mathbf{n}} V_{\mathbf{n}} \{1 + e^{\pi i (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)}\} e^{2\pi i \mathbf{n} \cdot \mathbf{r}/a}.$$

For points in  $k$ -space not too close to a surface of discontinuity, we find that when  $E$  is expressed as in (15), by the expression  $E = \omega k^2 + E_2$ ,  $\omega = \hbar^2/2ma^2$ , then

$$E_2 = -\frac{4}{\omega} \sum_{\mathbf{n}} \frac{|V_{\mathbf{n}}|^2 \cos^2 \pi (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)/2}{|\mathbf{n}|^2 + 2\mathbf{n} \cdot \mathbf{k}}. \quad (19)$$

The surfaces of discontinuity are thus given by the equation

$$\mathbf{n} \cdot \mathbf{k} + \frac{1}{2}|\mathbf{n}|^2 = 0,$$

where  $\mathbf{n}$  may have any set of values for which  $\cos \pi (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)/2$  does not vanish. The first set of planes is of the type (110), the second set is of the type (200), etc. Now the volume in  $\mathbf{k}$  space occupied by the Fermi distribution of a monovalent body centred lattice is unity. The radius of the Fermi sphere for free electrons is thus  $(4\pi/3)^{-1/3}$ , or 0.62. The smallest distance in  $\mathbf{k}$  space from the origin to a (110) plane is  $2^{-1/2}$ , or 0.71. The ratio of these two numbers is 0.87, while in the case of a simple cubic lattice the corresponding ratio is 0.985. In the body-centred lattice the Fermi surface, at least for free electrons, is hence relatively far from a surface of discontinuity.

A better approximation to the Fermi surface may readily be obtained. For points sufficiently far from a discontinuous surface, equation (19) may be used to calculate the eigenvalues  $E_{\mathbf{k}}$ . For points close to a single surface of discontinuity, the derivative of  $E_{\mathbf{k}}$  normal to the surface is dependent principally upon a single Fourier coefficient in the potential energy, and so is easily calculated. The results of such a calculation are shown in fig. 1 for the particular case of lithium.

Here we find Fourier coefficient  $|V_{110}|$  to be 0.79 electron volts.\* From this figure, or better from fig. 2, we may deduce that the Fermi surface differs but slightly from a sphere.

\* In calculating  $|V_{110}|$  we have taken the atomic wave functions for the free atom of Guillemin and Zener, 'Z. Physik,' vol. 61, p. 199 (1930). This is justifiable, since the valence electron is so spread out that  $|V_{110}|$  would be affected by less than 1% if the radius of its wave function were increased to infinity.

The above considerations show that little error is introduced by integrating over the spherical Fermi surface of free electrons. The details of the calculation will not be given. The eigenvalue  $E_{\mathbf{k}}$  was taken to be of the form (15), (19), and powers of  $E_{\mathbf{k}}$  greater than the second were neglected. All Fourier

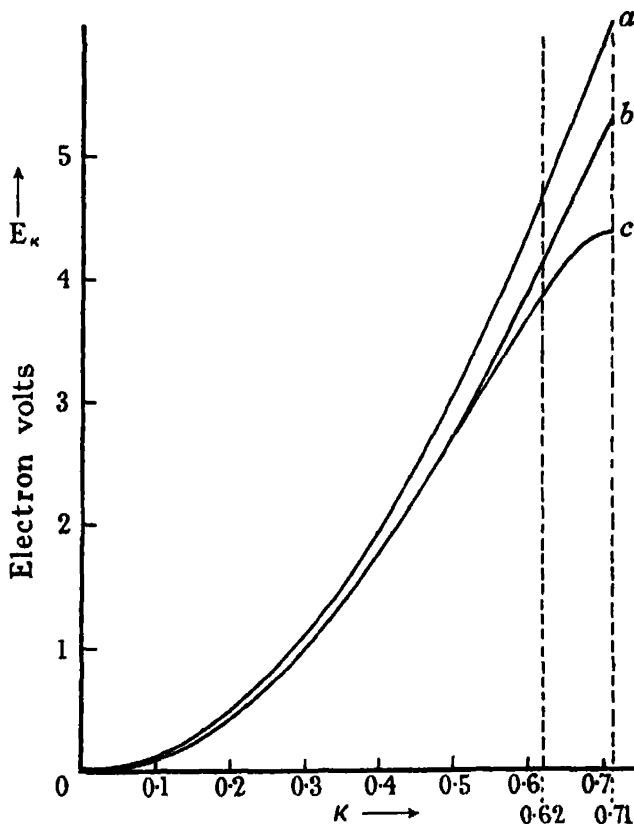


FIG. 1.—Eigenvalues for lithium. (a) gives the eigenvalues for free electrons. In this case the radius of the Fermi distribution is at  $|\mathbf{k}| = 0.62$ . (b) and (c) give the eigenvalues calculated by means of the actual potential of a Li lattice, the first along a direction of least perturbation ( $\eta = \zeta = 0$ ), the second along a direction of greatest perturbation ( $\xi = \eta$ ,  $\zeta = 0$ ). A discontinuity in the eigenvalue occurs along the latter direction at  $|\mathbf{k}| = 0.71$ .

coefficients except those of the type (110) were neglected. This is justified since the distance of the spherical surface to a plane of the type (200) is 1.38, as compared to 0.09 for the distance to a (110) plane. The integrals of the cross products of the terms in (19) were found, by numerical integration, to be negligible, amounting to less than 0.01 of the values of the integrals of the



square terms. The results are best expressed in terms of  $\epsilon$ , the energy at the surface of the Fermi distribution in the absence of a periodic potential, and  $\Delta E$ , the energy discontinuity across a (110) plane, i.e.,  $4|V_{110}|$ . The results are :

$$B_t = 1.16 (a^3/2ceR)^2 (\Delta E/\epsilon)^4$$

$$B_l = 1.23 (a^3/2ceR)^2 (\Delta E/\epsilon)^4$$

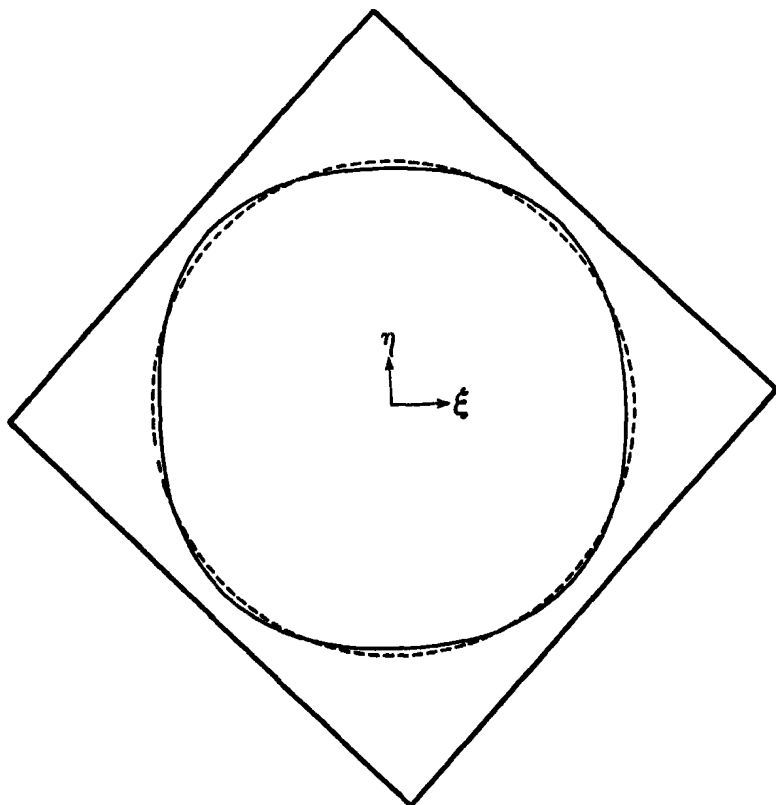


FIG. 2.—Intersection of the Fermi surface with the  $\zeta = 0$  plane. The surface for free electrons is given by dotted circle. The square represents the intersection of  $\zeta = 0$  plane with the (110) planes.

in the transverse and longitudinal case, respectively. The ratio  $B_t/B_l$  thus has the magnitude of unity, in agreement with experiments on all metals. In the particular case of lithium,  $N = 2/a^3 = 0.46 \times 10^{23}$ ,  $R = 1.3 \times 10^{-18}$  e.s.u. at liquid nitrogen temperature, the temperature of Kapitza's experiments,  $\epsilon = 4.6$  electron volts, and  $|V_{110}| = 0.79$  electron volts, as mentioned above. We obtain

$$B_t = 0.29 \times 10^{-18}.$$

The experimental value  $2.1 \times 10^{-12}$  is given by Kapitza\* as representing only the order of magnitude of  $B$ . This experimental value is approached by considering two refinements. The error made in neglecting all powers of  $E$ , greater than the second can be partially compensated by using the average value of  $dE/d|k|$ , as deduced from fig. 1, in place of the value for free electrons. This introduces a factor of about 2. The coefficient  $B$  was found to vary as the inverse fourth power of the distance from the Fermi surface to a (110) plane. Since the main contribution to the integrals comes from the regions closest to these planes, the error made in integrating over a spherical surface can be partially compensated by taking the radius of the sphere to have the same value as the magnitude of the radius vector to the actual surface nearest a (110) plane. This value is 0.64, in place of 0.62 for the radius of the sphere of unit volume. This final correction introduces a factor of 2.7, and brings the order of magnitude of the calculated and experimental value of  $B$  into complete agreement.

One of us (H. J.) is indebted to the Department of Scientific and Industrial Research for a maintenance grant.

#### 4. *Summary.*

The theory of the change of resistance of a metal in a magnetic field is developed for an arbitrary lattice, with the single assumption of a constant time between collisions for electrons at all points on the surface of the Fermi distribution. The theory is applied to Li, with excellent agreement with experiment.

\* 'Proc. Roy. Soc.,' A, vol. 123, p. 292 (1929).

---

## *The Structure of the Molecule of Nitrogen Dioxide from a Study of its Infra-Red Absorption Spectrum.*

By G. B. B. M. SUTHERLAND, Trinity College, Cambridge.

(Communicated by T. M. Lowry, F.R.S.—Received January 16, 1934.)

### *Introduction.*

In a previous paper,\* which gave an account of work on the infra-red absorption spectrum of nitrogen tetroxide there was necessarily included a preliminary survey of the spectrum of the dioxide, but no detailed description was given of the absorption bands, nor was any attempt made to deduce the structure of this molecule. Since that paper was communicated several other investigators have published accounts of work† on the infra-red and electronic spectra of these two molecules and conflicting views have been expressed regarding the structure of the dioxide molecule. Bailey and Cassie (*loc. cit.*) have favoured a linear symmetrical structure resembling that of carbon dioxide or carbon disulphide, while Harris, Benedict and King and also Schaffert (*loc. cit.*) have contended that the form of the molecule is triangular with the nitrogen atom at the apex of an isosceles triangle. None of these workers report having examined any of the infra-red bands under high dispersion, although the contours of two of the strongest of the dioxide bands were reported by Bailey and Cassie. Yet a detailed knowledge of any possible fine structure in the strongest bands and of the contours of at least some of the weaker bands is of supreme importance in deciding the structure of the molecule. The purpose of the present paper is to give an account of the examination of the infra-red spectrum of the dioxide under high dispersion and from a critical discussion of the present data to show that all the evidence is in favour of this molecule having a triangular form.

### *Experimental.*

For details of the experimental procedure the reader is referred to the earlier paper‡ on nitrogen tetroxide and Table I contains a list of the observed bands

\* Sutherland, 'Proc. Roy. Soc.,' A, vol. 141, p. 342 (1933), where references will also be found to earlier work on this molecule.

† Bailey and Cassie, 'Nature,' vol. 131, p. 239 (1933); Harris, Benedict and King, 'Nature,' vol. 131, p. 621 (1933); Bailey and Cassie, 'Nature,' vol. 131, p. 910 (1933); Schaffert, 'Nature,' vol. 131, p. 910 (1933); Schaffert, 'J. Chem. Phys.,' vol. 1, p. 507 (1933).

‡ Sutherland, *loc. cit.*

as there given. Since the examination of the absorption beyond  $1450\text{ cm.}^{-1}$  was done using a cell with rock-salt windows, it was not possible at that time to decide whether the absorption near  $1350\text{ cm.}^{-1}$  was a true absorption band of  $\text{NO}_2$  or whether it was due to the sodium nitrite which formed on the windows of the cell. The work of Schaffert (*loc. cit.*) in which a cell with very thin mica windows was used proves beyond doubt that there is an absorption band of  $\text{NO}_2$  at  $1373\text{ cm.}^{-1}$ . Again our observations were confined to wave-lengths less than  $14\text{ }\mu$  so we did not observe the absorption, which has also been confirmed by Schaffert (*loc. cit.*), reported by Bailey and Cassie (*loc. cit.*) at  $15.6\text{ }\mu$  ( $641\text{ cm.}^{-1}$ ).

Table I. - Infra-red Absorption Bands of Nitrogen Dioxide.

Frequency in $\text{cm.}^{-1}$ .	Intensity.	Frequency in $\text{cm.}^{-1}$ .	Intensity.
1000	Very weak	2910	Medium
1350 (?)	Weak	3120	Medium
1615	Very strong	3240	Weak
2220	Weak		

Of the bands in Table I the very intense one at  $1615\text{ cm.}^{-1}$  was also observed by Bailey and Cassie and by Schaffert (the observations of Harris, Benedict and King being confined to the region between  $1\text{ }\mu$  and  $4\text{ }\mu$ ) and was definitely associated by each of them with the  $\text{NO}_2$  molecule; while the band at  $2915\text{ cm.}^{-1}$  has been corroborated by Schaffert, by Harris, Benedict and King, and by the earlier work of Warburg and Leithauser,\* but was apparently missed by Bailey and Cassie. None of the other workers appear to have employed sufficient gas to pick up the weaker bands at  $1000\text{ cm.}^{-1}$ ,  $2220\text{ cm.}^{-1}$ ,  $3120\text{ cm.}^{-1}$  and  $3240\text{ cm.}^{-1}$ , although a knowledge of the definite presence or absence of combination bands is of great importance in determining the structure of the molecule. It has been found, however, on repeated examination of the bands in the neighbourhood of  $3\text{ }\mu$  that at room temperatures the band at  $3240\text{ cm.}^{-1}$  is extremely weak even for large concentrations of the gas, and although it apparently becomes stronger at higher temperatures there is a definite probability that this band is spurious and due to some other compound which appears on heating.

Harris, Benedict and King have suggested that there is a fundamental frequency of  $\text{NO}_2$  at  $752\text{ cm.}^{-1}$ , but our results on this point confirm those of Bailey and Cassie and of Schaffert, viz., that the absorption in this region is

\* Warburg and Leithauser, 'Ann. Physik,' vol. 23, p. 209 (1907); vol. 28, p. 313 (1909).

due solely to the  $\text{N}_2\text{O}_4$  molecule. In Table II we have accordingly assembled from the work of all the observers only those absorption bands which may legitimately be ascribed to the molecule of nitrogen dioxide and in all future discussion shall regard them as constituting the infra-red absorption spectrum of the molecule.

Table II.—The Infra-red Absorption Spectrum of  $\text{NO}_2$ .

Observer.	Frequency $\text{cm}^{-1}$ .	Wave-length $\mu$ .	Intensity.
Bailey and Cassie ; Schaffert	641	15.6	Strong
Sutherland	1000	10.0	Very weak
Schaffert	1373	7.28	Medium
All	1615	6.2	Very strong
Sutherland	2226	4.5	Medium
All except Bailey and Cassie	2910	3.44	Medium
Sutherland	3120	3.21	Medium
Sutherland	3240	3.09	Very weak (?)

### *Examination of the Bands.*

The normal procedure in the investigation of the infra-red spectrum of a substance is to make a preliminary investigation of it over as wide a region as possible, under low dispersion, with a prism spectrometer and then to pick out the regions of most intense absorption (presumably corresponding to the fundamental absorption frequencies) and examine these under high dispersion with a grating spectrometer. The information obtained in this way, when combined with a knowledge of the Raman spectrum of the molecule, is generally sufficient to enable one to determine uniquely the form and fundamental frequencies of the absorbing molecule, and the further examination of the weaker (presumably overtone or combination) bands under high dispersion is not as a rule of so much profit.\* However, where the regions of fundamental absorption are not readily accessible for high dispersion it becomes highly important to know something of the structure of the weaker combination bands and indeed the form of the molecule may be as uniquely established from an examination of a few combination bands as from a knowledge of the fundamentals now that the selection rules for the former have been worked out by Dennison.†

\* Except when carried out in the photographic region (*cf.* the work of Mecke and others) where it allows much greater accuracy in the determination of molecular constants, *e.g.*, moments of inertia and anharmonic constants.

† 'Rev. Mod. Phys.,' vol. 3, p. 280 (1931).

In the present instance there was little doubt that the very intense band at  $1615\text{ cm.}^{-1}$  was one of the fundamental frequencies and that another was to be identified with the band at  $641\text{ cm.}^{-1}$  observed by Bailey and Cassie. If the linear structure were correct then these should be the only fundamentals active in absorption, but if the triangular form obtained then the third fundamental ought also to be active and might be expected to lie\* somewhere near  $1300\text{ cm.}^{-1}$ . The work of Schaffert in establishing the presence of a band at  $1373\text{ cm.}^{-1}$  is then obviously of the greatest importance, since it is difficult to interpret this band as other than a fundamental frequency of the  $\text{NO}_2$  molecule. But the presence of this third fundamental can also be deduced solely from a consideration of the bands at  $2220\text{ cm.}^{-1}$  and  $2910\text{ cm.}^{-1}$  and it was those bands together with the intense one at  $1615\text{ cm.}^{-1}$  which were examined under high dispersion in the present investigation.

*The Band at  $1615\text{ cm.}^{-1}$ .*—Under low dispersion this band showed signs of a "doublet structure," a fact which Bailey and Cassie (*loc. cit.*) have also reported their estimate of the doublet separation being  $32\text{ cm.}^{-1}$ . The examination under high dispersion proved very difficult because of the background due to the very intense water vapour absorption in this region. Two different spectrometers were employed in the examination of this band † In the former was a grating of 1200 lines to the inch of which the second order was used, thus making it possible to resolve lines only  $1\text{ cm.}^{-1}$  apart, while in the latter was a grating of 4800 lines to the inch which when used in the first order was capable of resolving lines  $0.5\text{ cm.}^{-1}$  apart. With neither instrument was any sign of regular rotational fine structure observed in the band, the variations in absorption in certain parts of the band being entirely due to the very intense water vapour lines, which made observations often difficult and sometimes uncertain. It might be contended that the water vapour background obscured any rotational fine structure, but this was proved to be false by choosing small regions of the band in which there was no appreciable background and examining them with considerable care.

The absence of rotational fine structure with a spacing greater than  $0.5\text{ cm.}^{-1}$  is sufficient evidence in itself against a linear symmetrical form for this molecule. If such were so we can assume that the moment of inertia would be of

\* Cf. the "group frequencies" in  $\text{N}_2\text{O}_4$  of  $1265\text{ cm.}^{-1}$  and  $1360\text{ cm.}^{-1}$ , Sutherland, *loc. cit.*

† The author wishes to express his indebtedness to Professor E. F. Barker of the University of Michigan, U.S.A., for permission to use these instruments in the examination of this and other bands.

the same order of magnitude as that in the linear  $\text{N}_2\text{O}$  molecule. But the spacing of the lines in the  $\text{N}_2\text{O}$  spectrum\* is  $0.85 \text{ cm.}^{-1}$ , and so in a linear symmetrical  $\text{NO}_2$  molecule, because of the zero nuclear spin of oxygen, alternate rotation lines would be missing and the spacing constant would be of the order of  $1.5 \text{ cm.}^{-1}$ . On the other hand, a triangular molecule being an asymmetrical rotator would be expected to have a very complex rotational fine structure which would almost certainly be beyond the resolving powers of the spectrometers used. Thus the examination of this single band under high dispersion yields fairly conclusive evidence for a non-linear structure. It was extremely difficult on account of the water vapour lines to obtain anything like a clear cut contour of this band (and so no diagram is reproduced) but it did not seem to possess anything that might be interpreted as a Q branch, and our estimate of the separation between the two maxima is  $31 \text{ cm.}^{-1}$ .

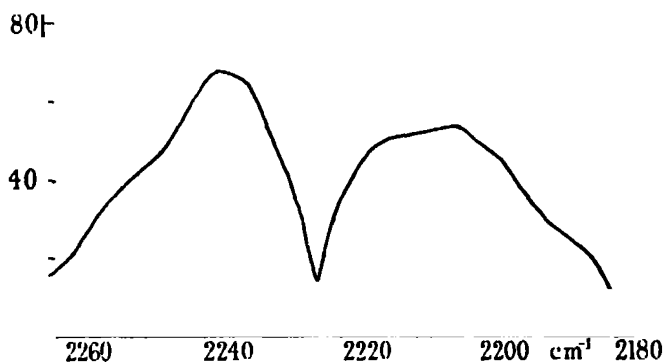


FIG. 1.

*The Band at  $2226 \text{ cm.}^{-1}$ .*—This band was examined using the first order of a grating with 4800 lines to the inch on a spectrometer capable of separating lines less than  $1 \text{ cm.}^{-1}$  apart. No sign of rotational fine structure was observed, but the band had a very pronounced gap in the centre indicating the complete absence of anything in the nature of a Q branch. It is important to notice that the mere presence of this band is an argument against the linear symmetrical structure, since it is obviously the combination of the fundamentals at  $641 \text{ cm.}^{-1}$  and  $1615 \text{ cm.}^{-1}$ , a combination which would be forbidden by the selection rules for combination bands in this type of molecule.† The “doublet separation” is approximately  $31 \text{ cm.}^{-1}$ . It is worth remarking that this band lies very close to one of the  $\text{N}_2\text{O}$  fundamentals, but after a careful consideration

\* Plyler and Barker, ‘Phys. Rev.’ vol. 38, p. 1827 (1931).

† Dennison, ‘Rev. Mod. Phys.’ loc. cit.

of the possibility of  $\text{N}_2\text{O}$  being present as an impurity we can definitely state that the probability of this is extremely small.

*The Band at 2910  $\text{cm}^{-1}$ .* Fig. 2.—This band was first examined on a spectrometer with a 7200 line grating (used in the first order), and secondly with a better spectrometer, in the first order of a 4800 line grating. In each case the resolving power was such that lines less than  $1 \text{ cm}^{-1}$  apart should have been separated with ease, and with neither instrument was any sign of a regular rotational fine structure observed. As the figure shows the band has

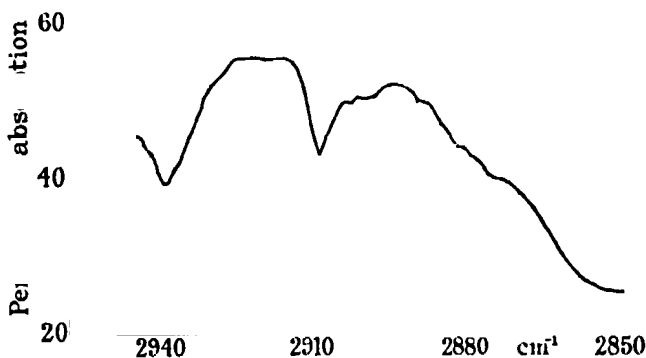


FIG. 2.

roughly a doublet structure with no indication of a Q branch, the separation between the maxima being approximately  $30 \text{ cm}^{-1}$ .

*The Bands at 1000  $\text{cm}^{-1}$ , 3120  $\text{cm}^{-1}$ , and 3240  $\text{cm}^{-1}$ .*—None of these bands was investigated in detail as the evidence collected from the others is quite sufficient to establish the structure of the molecule beyond any question or doubt. A rather cursory examination of the band at  $3120 \text{ cm}^{-1}$  did not reveal any structure whatever, the absorption rising to a maximum and dropping off again quite regularly. This might be taken to indicate the presence of a Q branch but the examination was not sufficiently detailed to justify such a conclusion. We have already remarked that there is some doubt whether the band at  $3240 \text{ cm}^{-1}$  is to be attributed to  $\text{NO}_2$  or to some impurity formed on heating the cell.

#### *Interpretation of the Bands and the Form of the Molecule.*

It is sufficiently clear from the evidence in the last section that the molecule is not linear and that the most probable values for the fundamental frequencies are  $641 \text{ cm}^{-1}$ ,  $1373 \text{ cm}^{-1}$  and  $1615 \text{ cm}^{-1}$ . We note that by such a choice of fundamentals we may account very simply for all of the other bands except the doubtful one at  $3240 \text{ cm}^{-1}$ ; thus the bands at  $2226 \text{ cm}^{-1}$  and  $1000 \text{ cm}^{-1}$



can be taken respectively as the sum and difference of the fundamentals at  $1615\text{ cm.}^{-1}$  and  $641\text{ cm.}^{-1}$ , while those at  $2910\text{ cm.}^{-1}$  and  $3120\text{ cm.}^{-1}$  can be interpreted respectively as the sum of  $1615\text{ cm.}^{-1}$  and  $1373\text{ cm.}^{-1}$  and the first overtone of  $1615\text{ cm.}^{-1}$  (Table III). In a triangular molecule of this type all possible overtone and combination bands are active and so by using still more gas it ought to be possible to detect additional absorption bands of  $\text{NO}_2$  at approximately  $1250\text{ cm.}^{-1}$ ,  $1900\text{ cm.}^{-1}$ ,  $2000\text{ cm.}^{-1}$ ,  $2700\text{ cm.}^{-1}$ , and so on; it is to be hoped that these will be verified in any further investigations on the spectrum of this molecule.\*

So far we have merely shown that the  $\text{NO}_2$  molecule must be triangular without being able to state anything regarding the exact form of the triangle. Our knowledge of the contours of the bands at  $2226\text{ cm.}^{-1}$ , and  $2910\text{ cm.}^{-1}$  enables us, however, to make certain deductions regarding the orientations of the axes of inertia when we make use of the selection rules of Dennison (*loc. cit.*) for the character of the combination bands in this type of molecule. These are :—

(1) that for  $n_3$  odd the electric moment vibrates perpendicular to the symmetry axis,

(2) that for  $n_3$  even the electric moment vibrates parallel to the symmetry axis,

(3) that a band arising from a vibration along the least axis of inertia should possess a Q branch,

(4) that a band arising from a vibration of the electric moment along the middle axis of inertia should possess no Q branch, where any absorption band is given by

$$\nu = n_1\nu_1 + n_2\nu_2 + n_3\nu_3,$$

and the characters of the vibrations denoted by  $\nu_1$ ,  $\nu_2$  and  $\nu_3$  are indicated in fig. 3. Our first problem is then to assign the fundamental frequencies at  $641\text{ cm.}^{-1}$ ,  $1373\text{ cm.}^{-1}$ , and  $1615\text{ cm.}^{-1}$  correctly among the three normal modes of vibration  $\nu_1$ ,  $\nu_2$  and  $\nu_3$ . From the nature of the vibration and from analogy with other triangular molecules (such as  $\text{SO}_2$  and  $\text{H}_2\text{O}$ ) it is quite certain that  $\nu_2$  is to be associated with the low frequency vibration at  $641\text{ cm.}^{-1}$ , but there is no simple criterion for differentiating  $\nu_1$  from  $\nu_3$ . We notice, however, from the first two of the selection rules that it is the value of the quantum number  $n_3$  which really determines the character of any combination

\* Some of these have already been detected by Harris and King, 'J. Chem., Phys.,' vol. 2, p. 51 (1934).

band, e.g., the bands  $\nu_3 + \nu_1$  and  $\nu_3 + \nu_2$  will have the same character as  $\nu_3$  itself. But we found that the bands at  $2220 \text{ cm.}^{-1}$  and  $2910 \text{ cm.}^{-1}$  were similar in that they each possessed no Q branch and therefore if we choose  $\nu_3 = 1615 \text{ cm.}^{-1}$  and  $\nu_1 = 1373 \text{ cm.}^{-1}$  the characters of these bands will be in accordance with the selection rules whereas if we interchange the assignments of  $\nu_1$  and  $\nu_3$  then the bands at  $2220 \text{ cm.}^{-1}$  and  $2910 \text{ cm.}^{-1}$  being respectively  $\nu_1 + \nu_2$  and  $\nu_1 + \nu_3$  ought to exhibit different characters. The fact that these two bands have been examined under such high dispersion as to leave no doubt whatever as to the similarity of their contours, makes the assignment of  $\nu_1 = 1373 \text{ cm.}^{-1}$  and  $\nu_3 = 1615 \text{ cm.}^{-1}$  unique.

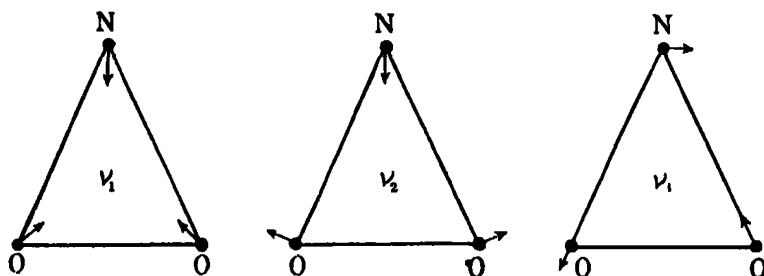


FIG. 3.

Once this has been decided, the form of the molecule follows immediately, for since  $\nu_3$  exhibits no Q branch it must arise from a vibration along the middle axis of inertia, while  $\nu_1$  and  $\nu_2$  will each be vibrations along the least axis. The fact that the moment of inertia about the symmetry axis is less than that about a perpendicular axis gives  $58^\circ$  as an upper limit for the ONO angle, and it seems probable from a consideration of the electronic structure (see next section) that the lower limit is about  $55^\circ$ . Table III illustrates the agreement between the observations and the theoretical consequences of the above interpretation. It will be observed that there is one serious inconsistency in the table, namely, that the band at  $641 \text{ cm.}^{-1}$  as  $\nu_2$  ought to exhibit a Q branch, whereas Bailey and Cassie observed it to be a doublet. This band ought to be re-examined under high dispersion,\* for it frequently happens that a Q branch is undetected when only low dispersion is employed. The  $\nu_1$  frequency at  $1373 \text{ cm.}^{-1}$  should, of course, also possess a Q branch and although Schaffert's curves (*loc. cit.*) would indicate that such is the case this band ought also to be re-examined to confirm it.

\* Professor Randall, of the University of Michigan, informs me that he intends to undertake this re-examination at an early date.

Table III.—Interpretation of the Infra-red Spectrum of  $\text{NO}_2$ .

Frequency observed cm. <sup>-1</sup> .	Assignment.	Frequency calculated cm. <sup>-1</sup> .	Character observed.	Character predicted.
641	$\nu_1$	—	Doublet	Q Branch
1000	$\nu_2 - \nu_3$	974	?	Doublet
1373	$\nu_1$	—	Probably Q Branch	Q Branch
1615	$\nu_3$	—	Doublet	Doublet
2220	$\nu_2 + \nu_3$	2256	Doublet	Doublet
2910	$\nu_2 + \nu_1$	2988	Doublet	Doublet
3120	$2\nu_3$	3230	Q Branch ?	Q Branch
3240	?	—	?	—

*Electronic Structure.*

It has been suggested by Pauling\* that there is a three electron bond in nitrogen dioxide and that the structure is to be represented thus :  $\ddot{\text{O}} : \ddot{\text{N}} :: \ddot{\text{O}} :$  where the oxygen atoms are not to be regarded as non-equivalent since (according to his concept of resonance between neighbouring bonds) the double bond regularly changes places with the single and the three electron bonds. Pauling has suggested this three electron bond in order to account for the relative saturation of the molecule and its small tendency to polymerize, and rules out the ring structure from consideration of the electric moment of the  $\text{NO}_2$  group in organic compounds and from an X-ray determination of the  $\text{ONO}$  angle as  $130^\circ$  in  $\text{NO}_2$  group in inorganic compounds. Now although the evidence in favour of the  $\text{NO}_2$  group having an obtuse-angled configuration in combination is fairly conclusive, it does not follow that the  $\text{NO}_2$  molecule must be similar and indeed we have seen that the infra-red data demand an  $\text{NO}_2$  molecule in which the  $\text{O}-\text{O}$  distance is less than the  $\text{N}-\text{O}$  distance. Such a condition must seem surprising at first sight but may arise somewhat as follows. The nitrogen atom has three  $2p$  electrons, the eigenfunctions of which will tend to adopt mutually perpendicular directions (the contribution of the shell of two  $2s$  electrons is to be neglected). Each oxygen atom (again neglecting the  $2s$  shell) has four  $2p$  electrons and we may imagine two of each of those coupled to zero, and the remaining two, each forming one  $pc$  bond with the nitrogen atom and one  $p$  bond with one another. This bond between the oxygen atoms will deform considerably the original  $90^\circ$  setting of the  $p$  eigenfunctions of the nitrogen and might easily reduce the  $\text{ONO}$  angle by 20 or 30 degrees. Since in

\* 'J. Amer. Chem. Soc.,' vol. 53, p. 3225 (1931).

fact the  $2p\sigma - 2p\sigma$  bond between the two oxygen atoms is nearly as strong as the  $2p\sigma - 2p\sigma$  bonds between the oxygen and nitrogen atoms one would expect the molecule to have the form of an equilateral triangle. It is a little difficult, however, when one takes account of the resonance between the  $2p\pi$  electrons to understand why the  $\text{ONO}$  angle is less than  $60^\circ$ . It may, of course, be that Dennison's selection rules regarding the contours of these bands are misleading when one is near the limiting case of a disc molecule (i.e., moments of inertia in the plane of the molecule equal). The whole question of the electronic structure of the  $\text{NO}_2$  molecule will require to be very fully investigated before these points can be decided. The author hopes to carry out such an investigation in collaboration with Dr. W. G. Penney in the near future. A discussion of the most probable values of the interatomic distances and forces is therefore of little significance at present and is reserved for a later paper.

The author is glad to acknowledge again the continued encouragement of Professor T. M. Lowry under whose guidance this work was begun. He wishes also to express his indebtedness to the Commonwealth Fund of New York for a Fellowship which enabled him to carry out the experimental part of the work in the Physics Laboratory of the University of Michigan, and to the Carnegie Trust for a Fellowship which has enabled him to continue the theoretical side of the investigation.

#### *Summary.*

(1) An account is given of new experimental work on the infra-red absorption spectrum of nitrogen dioxide under high dispersion.

(2) The present data on the infra-red absorption spectrum of this molecule are then critically analysed and shown to lead inevitably to the conclusion that the form of the molecule is triangular and not linear as has been suggested by some of the other observers.

(3) The fundamental frequencies of the molecule have the values  $641 \text{ cm.}^{-1}$ ,  $1373 \text{ cm.}^{-1}$  and  $1615 \text{ cm.}^{-1}$ .

(4) The contours of the infra-red bands indicate that the angle between the  $\text{NO}$  bonds is somewhat less than  $58^\circ$ , and a possible electronic configuration for such a ring structure is proposed.

---

*The Rate of Transformation in Aqueous Solution of Methylammonium Cyanate into Methylurea.*

By CHRISTINA C. MILLER, Carnegie Teaching Fellow, Chemistry Department,  
University of Edinburgh.

(Communicated by Sir James Walker, F.R.S.—Received January 18, 1934.

—Revised February 13, 1934.)

When two molecules A and B react irreversibly with formation of a third substance C, the velocity  $v$  of the reaction, according to the classical law of mass action, ought to be directly proportional to the product of the concentrations  $C_A$  and  $C_B$  of A and B at the time of observation, that is  $v = kC_AC_B$ , where  $k$  is a constant. Although this equation has been successfully applied to some reactions there are many exceptions, particularly reactions involving ions, where the irregularity is indicated by a progressive change in  $k$  with the initial concentration. Corrections have therefore had to be made in the above equation to account for the deviations from the classical mass action law. It has been suggested that the concentration terms should be replaced by thermodynamic activities, for the equilibrium constant  $K$  of the reaction  $A + B \rightleftharpoons C$  is given exactly by  $K = a_C/a_Aa_B$ , where  $a_A$ ,  $a_B$ , and  $a_C$  represent the activities of A, B, and C. Mere substitution of activities for concentrations in the velocity equation does not, however, explain the kinetic phenomena observed.

The views that up to the present have met with most success have been put forward by Brönsted.\* He supposes that probably all reactions involve the primary formation from the reactants of a fugitive complex which is in thermodynamic equilibrium with them. The rate of the reaction is determined by the rate of formation of this complex X which then breaks down spontaneously into the final product, at a rate proportional to the concentration of X. Brönsted finally deduces the following equation for the velocity of reaction :

$$v = ka_Aa_B/f_X = kC_AC_Bf_Af_B/f_X = kC_AC_BF,$$

where the  $f$  terms represent activity coefficients. The reaction kinetic factor  $F$  is in a great measure dependent on the ionic concentration of the solution, and, in general, it changes with the course of the reaction. The charge on X, in a reaction involving ions, is the algebraic sum of the charges on A and B, and its activity coefficient depends in the same way as that of a normal substance

\* 'Z. phys. Chem.,' vol. 102, p. 169 (1922).

on the medium. Brönsted's hypothesis has been tested in a great many reactions involving ions, particularly where the activity coefficients of the reactants are expected to undergo a considerable change, owing to medium changes, either due to continued reaction, or to the addition of various electrolytes.\*

Important bimolecular reactions to which the classical laws of chemical kinetics do not apply are the transformations in aqueous solution of ammonium and alkylammonium cyanates into the corresponding ureas. Walker and Hambly† concluded that the ions of the salts were the reacting molecules and that if dissociation were regarded as complete the velocity constants based on the law of mass action varied with the initial concentration of the cyanates. Walker and Kay‡ attributed the variations to incomplete dissociation of the cyanates and, following a procedure that was legitimate at that time, substituted for the stoichiometric concentrations the ionic concentrations as calculated by Arrhenius with the aid of the conductivity ratio  $\Lambda_C/\Lambda_0$ , where  $\Lambda_0$  and  $\Lambda_C$  represent equivalent conductivities at concentration  $C$  and at infinite dilution; that is,  $v = kC^2 (\Lambda_C/\Lambda_0)^2$ , when  $C_A = C_B = C$ . In this way velocity constants were obtained for ammonium cyanate which, although somewhat variable, varied irregularly about a mean and showed no definite trend as concentration steadily altered. Now the modern view is that dissociation of a uni-univalent salt like ammonium cyanate, in dilute aqueous solution, is virtually complete, therefore Walker and Kay's explanation of the deviation from the simple velocity equation  $v = kC_A C_B$  is inadequate, although the insertion of the conductivity ratio in the velocity equation is not necessarily wrong. The correct velocity equation is evidently of the type  $v = kC_A C_B F$ , where  $F$  according to Brönsted is  $f_A f_B / f_X$  and cannot at present be evaluated. Its magnitude according to Walker is satisfactorily represented by the square of the conductivity ratio, but there is no theoretical basis for its insertion in the reaction rate equation.

Although Walker and Kay's velocity measurements were of considerable accuracy various approximations were made in calculating the dependence of the rate on the supposed true concentrations of the ions that render it doubtful if the equation  $v = kC^2 (\Lambda_C/\Lambda_0)^2$  represents the experimental results quite as well as was believed. It was therefore proposed to re-examine the rate of transformation, and the equivalent conductivity in aqueous solution,

\* See La Mer, 'Chem. Rev.', vol. 10, p. 179 (1932).

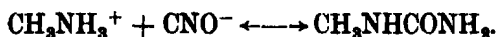
† 'J. Chem. Soc.', vol. 67, p. 746 (1895).

‡ 'J. Chem. Soc.', vol. 71, p. 489 (1897).

of a suitable cyanate, taking all precautions to obtain exact results. For this purpose methylammonium cyanate, the rate of change of which was previously examined by Walker and Appleyard,\* was selected. It could be prepared in a pure state, and its transformation was less disturbed by secondary reactions than were those of other alkylammonium cyanates and ammonium cyanate.

### Experimental.

The transformation of methylammonium cyanate into methylurea is a reversible bimolecular process which may be represented thus :



If complete dissociation is assumed, and if  $C$  represents the original concentrations of methylammonium and cyanate ions,  $B$  that of methylurea, and  $x$  the amount of methylammonium cyanate transformed after time  $t$ , then the rate of formation of methylurea, according to the concentration mass action law, is given by  $dx/dt = k(C - x)^2 - k_1(B + x)$ ,  $k$  and  $k_1$  being respectively the velocity constants of the forward and reverse reactions. When  $k$  is large compared with  $k_1$ , and  $B$  is small, the reverse reaction may be ignored and

$$k = x/tC(C - x). \quad (1)$$

When the initial concentrations of methylammonium and cyanate ions are  $C_A$  and  $C_B$  respectively, the corresponding equation is

$$k = \frac{2 \cdot 303}{(C_A - C_B)t} \times \log \frac{(C_A - x)C_B}{(C_B - x)C_A}. \quad (2)$$

If, on the other hand, the rate of reaction is proportional to  $F(C - x)^2$ , where  $F$  is a factor that varies with concentration, the true velocity constant  $k'$  is given by

$$x/tFC(C - x), \quad (3)$$

for the range of concentration over which  $F$  may be considered constant. As this range is likely to be small it follows that velocity constants should be calculated for separate small portions of the total change. This method was adopted with methylammonium cyanate.

*The Quantitative Determination of Cyanates as Silver Cyanate.*—In the reaction velocity experiments the course of the change was followed by deter-

\* 'J. Chem. Soc.,' vol. 69, p. 193 (1896).

mining the cyanate initially and after appropriate intervals, according to a standardized procedure. 5, 10, or 25 gm. portions of the solution, depending on the concentration, were added to suitable silver nitrate solutions, to which excess of silver nitrate was then added. After about an hour at 18° C. the precipitates were filtered off and washed at 18° with the minimum amount of a dilute potassium nitrate solution, and dissolved in dilute nitric acid for the determination of silver with a standard solution of ammonium thiocyanate. An iron alum indicator was employed and the end-points were fixed colorimetrically. The ammonium thiocyanate solution was standardized with silver cyanate containing the theoretical amount of silver. As solubility errors could not be neglected, particularly in dilute cyanate solutions, the cyanate in standard solutions of carbonate-free potassium cyanate was determined in exactly the same way. From the differences between the theoretical and the experimental results, solubility corrections for any desired cyanate concentrations were derived, and subsequently applied in *all* the determinations of methylammonium cyanate. Carbonate-contaminated potassium cyanate was purified as follows. To a concentrated aqueous solution pure acetone was added until the aqueous layer was negligible in size, and the bulk of the carbonate was precipitated along with potassium cyanate. The acetone layer was removed and dissolved cyanate was precipitated by addition of dry acetone. Further purification was effected by dissolving the precipitate in the *minimum* amount of water, and reprecipitating by adding purified absolute ethyl alcohol. The crystals that separated were washed with a little alcohol and anhydrous ether and kept in an evacuated desiccator containing concentrated sulphuric acid to avoid decomposition.

*Preparation of Methylammonium Cyanate and Methylurea.*—Solid methylammonium cyanate was prepared by very slowly mixing, at -18° C., a filtered ethereal solution of cyanic acid with an ethereal solution of methylamine, precautions being taken to exclude moisture. The precipitated methylammonium cyanate was rapidly filtered and transferred to a desiccator containing sulphuric acid; the desiccator was evacuated. As the solid could not be kept untransformed for any length of time, and could not be directly weighed, it had to be brought into solution very soon and weighed indirectly by adding it quickly, with stirring, to a known weight of water. More dilute solutions were then prepared, and either used immediately, or kept at 3° C. until required. Week old solutions were usually rejected. Freshly prepared solutions contained 98% of the theoretical amount of methylammonium cyanate, the remainder having been converted to methylurea.



Methylurea was prepared from methylammonium cyanate by allowing it to undergo spontaneous transformation at room temperature, preferably in a desiccator.

*Reaction Velocity Experiments.*—Experiments were made at about 40° C. A suitable quantity of the cyanate solution contained in a stoppered flask, was placed in a thermostat, the temperature of which during the course of an experiment varied  $\pm 0.01^\circ$  from the mean. After 20 minutes four similar portions of the solution were removed in rapid succession by means of pipettes and added to silver nitrate solutions, the times being noted. The average amount of silver cyanate subsequently found corresponded to the average time at which the cyanate solutions had been added to the silver nitrate solutions. After an interval of time involving a fall in concentration in the cyanate solution of about 0.01 M, at the higher initial concentrations, and less at the lower, the above procedure was repeated. If the lapse of time between the series of analyses was short, a further set of four analyses was made, but the total time of experiment never exceeded about 8 hours, and was generally much less. Density determinations were made in order to convert weight molarities to volume molarities.

Simultaneously with the above experiments others were carried out for the purpose of determining the equilibrium point of the reaction. Solutions of the cyanate, 0.01, 0.005, and 0.002 M, each containing sufficient methylurea to make it decimolar with respect to cyanate plus methylurea, were heated to 40° C. and analysed immediately and after 5 hours. From the observed rates of change the concentration at which the rate became nil was found to be 0.0003 M. (see p. 297). It was consequently unnecessary to consider the reverse reaction, even with low initial concentrations of cyanate, provided that the concentration of methylurea was kept low. Actually, the velocity experiments involving concentrations of cyanate above 0.03 M were performed in presence of methylurea to 0.1 M, but it did not influence the results.

A few experiments were made at 50° C. in a somewhat different manner, the results being less satisfactory than those at 40° C. The single velocity constant, obtained by averaging several results, and given below Table I, for the subsequent calculation of the temperature coefficient of the reaction, is regarded as reliable.

The results of the velocity experiments, all referred to the temperature 39.88° C., are given in Table I, and fig. 1.  $C$ ,  $B$ ,  $x$ , and  $t$  have the significance attached to them on p. 290, and  $C_{\text{aver}}$  represents the average concentration of the interval to which  $k$  refers. Concentrations are given in moles per litre,

times in minutes, and velocity constants, calculated from equation (1), in moles per litre per minute.

Table I.

<i>t.</i>	<i>B.</i>	<i>C.</i>	<i>x.</i>	<i>k.</i>	<i>Caver.</i>
60.6	0.0051	0.09523	0.00826	0.01646	0.0911
90.7	0.0130	0.08732	0.01021	0.01672	0.0822
121.2	0.0232	0.07711	0.01068	0.01720	0.0718
131.1	0.0313	0.06898	0.00944	0.01754	0.0643
171.5	0.0408	0.05954	0.00928	0.01808	0.0549
212.0	0.0497	0.05056	0.00831	0.01835	0.0464
303.0	0.0588	0.04152	0.00804	0.01910	0.03750
359.9	0.0661	0.03191	0.00594	0.01992	0.02894
358.1	0.0011	0.03194	0.00593	0.01993	0.02898
360.0	0.0005	0.02189	0.00310	0.02094	0.02034
360.0	0.0003	0.02132	0.00294	0.02084	0.01985
481.2	0.0002	0.009740	0.000920	0.02225	0.00928
460.0	0.0004	0.009585	0.000845	0.02193	0.00916

$k_{50.00^\circ} = 0.0522$  at average concentration 0.0822 M.

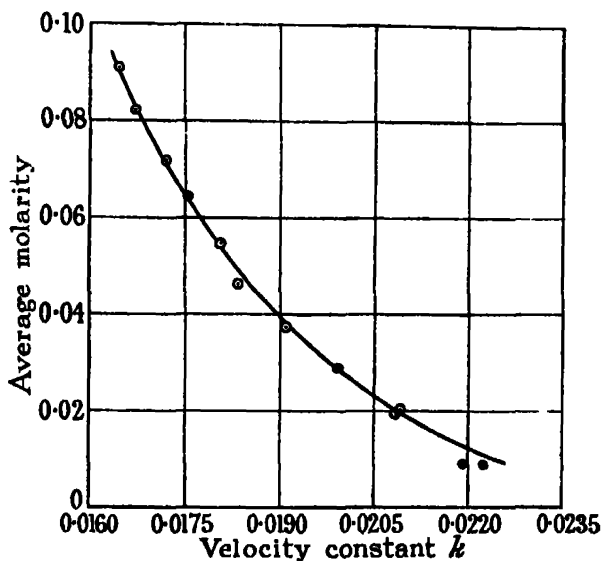


FIG. 1.

The increase in the value of  $k$  with increasing dilution is unmistakable, and is fairly regular. It was evident that the analytical method was becoming deficient in accuracy in the centinormal solutions, and that no useful purpose was to be served by examining more dilute solutions by this means.

*The Equivalent Conductivity of Methylammonium Cyanate.*—The equivalent conductivity of methylammonium cyanate was determined directly at 39.88° C. The equipment consisted of a rotating commutator, a galvanometer, a 3 metre drum-wound bridge wire, a box of standard resistances, and three conductivity cells with constants 0.75, 0.18, and 0.045 approximately. The cell constants of the first two cells were determined at 39.88° C. with the aid of H. and E. Parker's\* solutions of pure potassium chloride, containing 0.1 and 0.01 equivalents per cubic decimetre at 0° C. Their specific conductivities were calculated from the formula given for temperatures between 0° and 30° C. For the determination of the constant of the third cell, dilute solutions of potassium chloride, comparable with those of the methylammonium cyanate to be used in the cell, were employed at 25° C., and Jeffery and Vogel's† formula was accepted for the calculation of the specific conductivities. The cell constant at 40° C. was assumed to be the same as at 25° C. The specific conductivity of the water (about 2 gemmhos at 40° C.) was throughout subtracted from the specific conductivities observed.

In covering the concentration range 0.1 to 0.00025 M two methods of determining the composition of the methylammonium cyanate solutions, at the exact time of making the electrical measurements, were necessary. The first was suitable for concentrations 0.1–0.01 M. The solution in a stoppered flask was heated to 40° C. and part was transferred to the cell which was kept in the thermostat. After a short time the conductivity was measured, and, immediately afterwards, two or three analyses were made of the solution in the flask. From the difference between the average times of taking the bridge readings and starting the analyses, and the known rate of transformation of the cyanate at the concentration in question, the composition of the solution corresponding to the time of determining the conductivity was calculated, and thence the equivalent conductivity was found. The second method, which was suitable for concentrations below 0.0075 M, involved the analysis of a moderately concentrated solution at the time of diluting part of it to the strength required for immediate determination of the conductivity. To a known weight of water at 40° C. an appropriate amount was added from a weight burette; the cell was at once filled with the dilute solution and analyses of the concentrated solution were begun. Within 10 to 15 minutes the conductivity was determined, and again after an hour, the two measurements permitting the calculation of the conductivity at the time of preparing the dilute solution.

\* 'J. Amer. Chem. Soc.,' vol. 46, p. 312 (1924).

† 'J. Chem. Soc.,' p. 1715 (1931).

In Table II are given the values of the equivalent conductivities,  $\Lambda_c$ , corresponding to the various concentrations of cyanate,  $C$ , in moles per litre, and at the temperature  $39.88^\circ \text{C}$ . The concentration of methylurea seldom exceeded

Table II.

$C$ .	$\Lambda_c$ .	$C$ .	$\Lambda_c$ .
0.08610	131.5	0.00980	149.2
0.07686	133.5	0.005854	151.2
0.05936	136.5	0.002489	154.35
0.04636	138.8	0.000912	156.8
0.03901	140.2	0.0004965	157.1
0.02914	142.4	0.0002647	158.6
0.01920	144.0	0.0002391	158.6

10% of the cyanate concentration, and it was generally much less. The results are also shown graphically in fig. 2.

The last five results when substituted in Onsager's\* equation give for  $\Lambda_0$  the values 160.4, 160.4, 159.8, 160.6, 160.5—average 160.3.

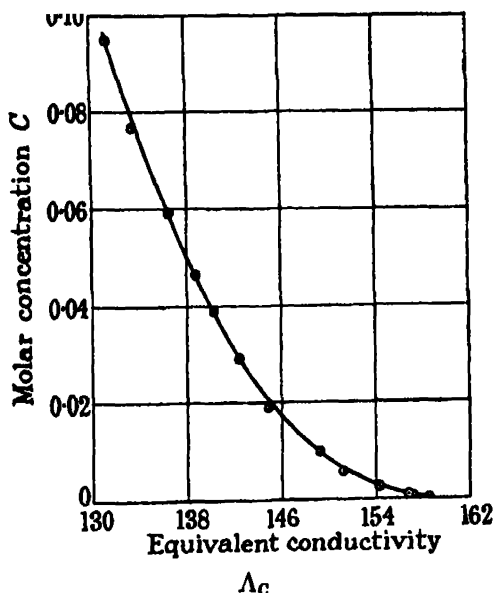


FIG. 2.

In order to find if  $(\Lambda_c/\Lambda_0)^2$  is the factor  $F$  required in the reaction velocity equation the velocity constants  $k$  in Table I were divided by the values of

\* Davies, "The Conductivity of Solutions," p. 35 (1933).

$(\Lambda_c/\Lambda_o)^2$  corresponding to the average concentrations to which values of  $k$  refer. The results are given in Table III. The duplicated velocity constants for concentrations 0.02 and 0.01 M have been averaged.

Table III.

No.	Caver.	$k$ .	$\Lambda_c$ .	$\Lambda_c/\Lambda_o$ .	$k' = k/(\Lambda_c/\Lambda_o)^2$ .
1	0.0911	0.01646	131.9	0.823	0.0243
2	0.0822	0.01672	133.0	0.830	0.0243
3	0.0718	0.01720	134.5	0.839	0.0244
4	0.0643	0.01754	135.65	0.846	0.0245
5	0.0549	0.01808	137.25	0.856	0.0247
6	0.0464	0.01835	138.8	0.866	0.0245
7	0.0375	0.01910	140.5	0.876	0.0249
8	0.0290	0.01993	142.35	0.888	0.0253
9	0.0201	0.02089	144.75	0.903	0.0256
10	0.00922	0.02209	149.45	0.933	0.0254

Although there is some irregularity, the figures in the last column show an unmistakable rise as concentration falls, thus indicating that insertion of  $(\Lambda_c/\Lambda_o)^2$  in the velocity equation does not lead to a true velocity constant for all concentrations, as was suggested by Walker and Kay (*loc. cit.*).

*The Velocity Constant at Infinite Dilution.*—If the factor  $F$  in the reaction velocity equation (3) is given by  $f_A f_B$  or  $f_A f_B / f_X$ , it follows that in infinitely dilute solution, where activity coefficients are equal to unity,  $F = 1$ , and velocity constants derived from equation (1) reach a limiting value  $k_0$ , which is simply  $k'$ . If  $k_0$  is known the factor  $F$  can be calculated for higher concentrations. An attempt was made to determine velocity constants in very dilute solutions of methylammonium cyanate at 40° C. by observing the conductivity change in 4 hours. Unexpected results were obtained. The conductivity of a 0.0025 M solution remained practically steady, although careful analysis revealed the disappearance of cyanate at more than twice the rate anticipated. In solutions more dilute conductivity increased at a gradually diminishing rate and eventually decreased slowly, and in solutions more concentrated the rate of conductivity decrease was less than expected. These results were accounted for by the formation of methylammonium ammonium carbonate, the quantity of which, owing to solubility effects, was insufficient to cause silver carbonate to remain associated with precipitated silver cyanate after filtration and washing. Its conductivity was, however, double that of the cyanate from which it was produced. A stationary conductivity could be explained by the simultaneous conversion of the same amounts of cyanate

to carbonate and to methylurea. Conductivities that initially increased and eventually decreased were, no doubt, a consequence of the influence on the reaction of the cations of the carbonate.

It might be thought that all the results in Table I are seriously in error because of carbonate formation, but this is probably not so, because the ratio of the rate of carbonate formation to that of methylurea formation seemed to decrease very rapidly as concentration increased, so that the former effect became negligible at higher concentrations. The effect of carbonate formation on the determination of the equilibrium point of the reaction had also to be borne in mind. While the result shown on p. 292 was undoubtedly too low, the corrected value was probably not sufficiently great to render equation (1) invalid for the calculation of velocity constants. On account of experimental difficulties attempts to determine velocity constants directly at high dilutions were abandoned.

The limiting value of the velocity constant at infinite dilution was found by extrapolation. Use was made of the fact that, if  $k_1, \dots, k_{10}$  and  $\Lambda_1, \dots, \Lambda_{10}$  represent respectively the values of  $k$  and  $\Lambda$  corresponding to the numbers in the first column in Table III, the value of  $(\log k_1 - \log k_n)/(\Lambda_1 - \Lambda_n)$ , where  $n = 2, \dots, 10$ , tended to increase very slightly as  $n$  increased, and the corresponding concentration progressively diminished. On the assumption that the increase was regular the value of  $(\log k_1 - \log k_0)/(\Lambda_1 - \Lambda_0)$  was calculated to be 0.0079, which leads to  $k_0 = 0.0276$ .

Table IV.

$\mu$ .	$k$ .	$F$ .	$\gamma^2$ .
0.0924	0.01646	0.596	0.602
0.0833	0.01672	0.606	0.613
0.0727	0.01720	0.623	0.627
0.0650	0.01754	0.635	0.640
0.05552	0.01808	0.655	0.656
0.04688	0.01835	0.665	0.674
0.03787	0.01910	0.692	0.697
0.02923	0.01993	0.722	0.723
0.02027	0.02089	0.757	0.759
0.00930	0.02209	0.800	0.821

*Calculation of the Factor F.*—In Table IV are given the values of  $F$  calculated from the relation  $k/F = k_0 = 0.0276$ . In order to find if  $F$  bears any resemblance to a factor of the type  $f_A f_B / f_X$ , the values of the square of the mean stoichiometric activity coefficient,  $\gamma$ , of the typical uni-univalent electrolyte, potassium chloride, are given in the last column. The values of  $\gamma$  were

determined graphically, from Spencer's\* data at 25° C., for the molalities to which the velocity constants refer. Concentrations have been converted to molalities and are given under  $\mu$  in the first column.  $\mu$  is the ionic strength of the solutions.

The resemblance between  $F$  and  $\gamma^2$  is marked and suggests that, if the activity of methylammonium cyanate is similar to that of potassium chloride,  $F$  does contain the factors  $f_A$  and  $f_B$ . One cannot, of course, distinguish between  $F = f_A f_B / f_X$  and  $F = f_A f_B$  without knowing something about the true values of  $f_A$  and  $f_B$ .

In very dilute solution the value of  $\log f_A f_B / f_X$ , where  $f_X$  is assumed to be unity, is given approximately by  $-\sqrt{\mu}$ .† If  $\log k = \log k_0 + \log f_A f_B / f_X$  its value at high dilution ought to become  $\log k_0 - \sqrt{\mu}$ , and  $\log k$  plotted against  $\sqrt{\mu}$  should give a line of limiting slope  $-1$ . In fig. 3  $\log k + 2$  is plotted against  $\sqrt{\mu}$ . The curve has a slope in the right direction, tending towards a value of the correct magnitude. It bears a marked resemblance to that found

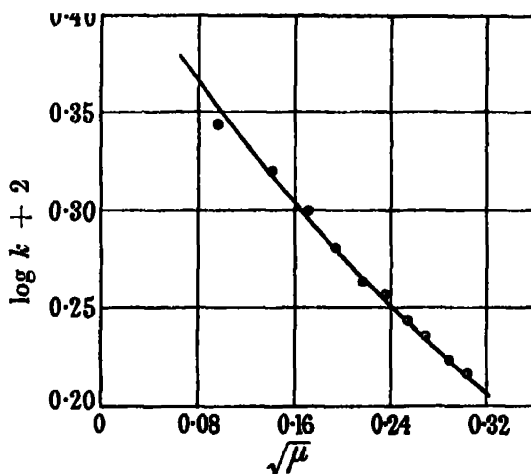


FIG. 3.

by Livingston\* for an entirely different reaction of the same ionic type, and for which the corresponding curve is most clearly given by La Mer (*loc. cit.*). The slope of the tangent to the curve at  $\sqrt{\mu} = 0.18$  is about 70% of the limiting value, and it is notable that the plots of  $\log \gamma^2$  for various uni-univalent

\* 'J. Amer. Chem. Soc.,' vol. 54, p. 4490 (1932).

† Glasstone, "Recent Advances in Physical Chemistry," p. 402 (1931).

‡ 'J. Amer. Chem. Soc.,' vol. 48, p. 53 (1926).

salts against  $\sqrt{\mu}$  exhibit similar or smaller deviations from the limiting value.\*

*The Effect of Neutral Salts on the Rate of Transformation.*—It was proposed to find how the factor  $F$  varied when the transformation of methylammonium cyanate proceeded in the presence of various neutral salts, some of which contained a common ion. The following salts were used, methylammonium nitrate, methylammonium sulphate, potassium, barium and magnesium cyanates and nitrates, potassium and magnesium sulphates, and barium isosuccinate. The first two salts were prepared from methylamine and the appropriate acids. A solution of magnesium cyanate was made by shaking a filtered ethereal solution of cyanic acid with solid magnesium carbonate and a little water until all of the acid had reacted. Ether was expelled, excess magnesium carbonate was filtered off, and the cyanate and carbonate contents of the filtrate were determined. Barium cyanate solution was prepared in similar fashion and was practically carbonate-free. Barium isosuccinate was made by boiling an aqueous solution of isosuccinic acid with barium carbonate, excess of which was removed in the cold. The solid that separated from the filtrate on evaporation was dried and analysed before use. An analysed solution of magnesium nitrate was employed, and the remaining salts were added in solid form to the cyanate solutions.

Reaction velocity experiments were carried out at 40° C. with solutions about 0.07 M with respect to methylammonium cyanate and 0.05 or 0.1 M with respect to the added salt. Duplicate analyses were made initially and finally, with suitable modifications when metallic cyanates were present. As solutions of metallic cyanates undergo slight decomposition, with production of carbonates and urea,† blank experiments were made by heating their solutions alone, and analysing them in the same way as the solutions containing methylammonium cyanate. Appropriate small corrections were made in calculating velocity constants, and the ionic strengths of the solutions.

In Table V,  $C$  represents the concentration of methylammonium cyanate and  $x$  the amount transformed in time  $t$ ;  $M$  is the molarity of the neutral salt. For metallic cyanates  $M$  is the average concentration, during an experiment, of the cyanate, exclusive of a small amount of carbonate. Velocity constants were calculated from equations (1) and (2),  $C_A$  and  $C_B$  in the latter being derived from  $C$  and  $M$ . All results refer to 39.88° C. The factors  $F$  required to make the velocity constants equal to 0.0276 are compared with the values

\* Noyes, 'J. Amer. Chem. Soc.,' vol. 46, p. 1098 (1924).

† O. and I. Masson, 'Z. phys. Chem.,' vol. 70, p. 290 (1910).



of  $\gamma^2$  (see pp. 297, 298) corresponding to the average ionic strengths  $\mu$  (molal) of the solutions.

It is evident that if dissociation is assumed complete and the ions are regarded as the reactive molecules there is no essential difference between the effects on the velocity constant of added salts with a common ion and those without. Increase in ionic strength is accompanied by a diminution in the value of the velocity constant involving concentration terms only. The same observation

Table V.

Added salt.	M.	<i>t</i> .	C.	<i>z</i> .	<i>k</i> .	$\mu$ .	F.	$\gamma^2$ .
CH <sub>3</sub> NH <sub>2</sub> · HNO <sub>3</sub>	0.0995	60.8	0.0702	0.0097	0.0148	0.168	0.535	0.535
KCNO	0.1002	80.1	0.0704	0.0097	0.0147	0.169	0.540	0.535
KNO <sub>3</sub> ...	0.0995	180.6	0.0684	0.0108	0.0152	0.165	0.550	0.535
(CH <sub>3</sub> NH <sub>2</sub> ) <sub>2</sub> · H <sub>2</sub> SO <sub>4</sub>	0.0498	61.4	0.0707	0.0091	0.0136	0.220	0.495	0.510
K <sub>2</sub> SO <sub>4</sub>	0.0505	180.8	0.0686	0.0100	0.0138	0.219	0.500	0.510
Ba(CNO) <sub>2</sub>	0.0501	80.6	0.0619	0.0083	0.0150	0.211	0.545	0.515
Ba(NO <sub>3</sub> ) <sub>2</sub>	0.0502	180.2	0.0614	0.0086	0.0147	0.211	0.535	0.515
Mg(CNO) <sub>2</sub>	0.0471	61.2	0.0665	0.0087	0.0147	0.213	0.535	0.510
Mg(NO <sub>3</sub> ) <sub>2</sub>	0.0495	180.4	0.0610	0.0084	0.0145	0.209	0.525	0.515
K <sub>2</sub> SO <sub>4</sub>	0.0992	180.3	0.0679	0.0088	0.0122	0.368	0.440	0.480
MgSO <sub>4</sub>	0.1003	180.2	0.0669	0.0086	0.0122	0.471	0.440	0.435
Ba isosuccinate	0.0965	180.5	0.0602	0.0077	0.0135	0.450	0.485	0.440

is made (Table I) when methylammonium cyanate is alone responsible for changes in ionic strength. Between the kinetic factor  $F$  and  $\gamma^2$  for potassium chloride there is considerable parallelism, which may indicate that kinetic activities and thermodynamic activities are not very different. The percentage differences between  $F$  and  $\gamma^2$  are comparatively small and vary somewhat irregularly from salt to salt. Indeed, if  $F$  includes the activity coefficients of methylammonium and cyanate ions the variation is distinctly less than one might expect by comparison with the known influence of many of the above salts on the activity coefficients of silver chloride, thallous iodate, etc.\* This suggests that  $F$  would be better represented by  $f_A f_B$  divided by a third factor that shows relative specific salt effects similar to those of  $f_A$  and  $f_B$ . Such a factor might be  $f_X$ , the activity coefficient of an uncharged critical complex. Although the above salts have a specific influence on the activity coefficients of uncharged molecules,† it is impossible to say whether the relative effects

\* Neuman, 'J. Amer. Chem. Soc.,' vol. 54, p. 2195 (1932); La Mer and Goldman, *ibid.*, vol. 51, p. 2632 (1929).

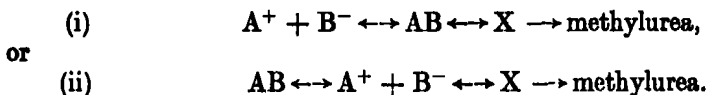
† Randall and Failey, 'Chem. Rev.,' vol. 4, p. 285 (1927); Larason, 'Z. phys. Chem.,' A, vol. 148, p. 304 (1930); vol. 153, p. 299 (1932).

of any two given solvent salts on the activity coefficients of ions A and B and on the activity coefficient of a critical complex would be similar. If they are, then  $f_A f_B / f_X$  better represents factor F than  $f_A f_B$ .\*

### *The Mechanism of the Reaction.*

The temperature coefficient of the reaction, calculated from the velocity constants 0.01672 at 39.88° C. and 0.0522 at 50.0° C., for the average molarity 0.082, is  $k_{50.0^\circ} / k_{40.0^\circ} = 3.08$ , and the critical increment  $E_A$  calculated from Arrhenius's equation  $k = Ze^{-\frac{E_A}{RT}}$ , where Z is a constant, is 22,570 calories.

While it is evident that the ions of methylammonium cyanate play an important part in its transformation to methylurea, it has yet to be shown how they act. If one considers the possibility of the cyanate (AB) being incompletely dissociated into its ions ( $A^+$  and  $B^-$ ), either of the following mechanisms involving formation of a critical complex X is feasible :



As the velocities of ionization and recombination are very great compared with the rate of methylurea formation the former can be neglected in comparison with the latter. In (i) the velocity of formation of methylurea according to Brönsted is

$$v = k_{\text{uni}} C_{AB} \cdot f_{AB} / f_X = k_{\text{uni}} / K \times C_A C_B f_A f_B / f_X,$$

and in (ii),

$$v = k_{\text{bi}} C_A C_B f_A f_B / f_X,$$

$k_{\text{uni}}$  and  $k_{\text{bi}}$  being unimolecular and bimolecular velocity constants respectively, and K the true dissociation constant of AB. Brönsted's equation for a reaction between two uni-univalent ions is therefore satisfied whether the ions or the unionized molecules are regarded as the final reactive molecules. With ammonium cyanate Chattaway† has objected to the ions being regarded as the reactive molecules, principally because the reaction proceeds much faster in alcoholic solutions, in which the number of ions is likely to be diminished. If mechanism (i) is correct, substitution of alcohol for water will

\* *Note*.—When this work was complete Warner and Stitt, 'J. Amer. Chem. Soc.,' vol. 55, p. 4807 (1933), published an account of their investigations on the kinetic salt effect in the transformation of ammonium cyanate to urea. Their results, although not of a high degree of accuracy, are in general agreement with those given above.

† 'J. Chem. Soc.,' vol. 101, p. 170 (1912).

bring about a change in activities, resulting in a readjustment of the equilibrium, whereby  $k_{uni}$  will remain constant and  $k_{bi}$  will vary inversely as the dissociation constants of the cyanate in the two solvents. As shown by Ross\* the values of  $k_{bi}$  for ammonium cyanate in alcohol and water at 32° C., corrected by means of the conductivity ratio, are 1.57 and 0.00595. These may be taken as an approximation to the values at infinite dilution; their ratio is 264. The dissociation constants of cyanates are unknown, but it is significant that K for silver nitrate in water at 25° C† has been calculated to be 400 times the value in alcohol at the same temperature.‡

The bimolecular mechanism involving the ions may next be considered, complete dissociation being assumed. Bimolecular reactions in all probability involve collisions between molecules that are activated and possess an energy in excess of the average E. It has been pointed out by Christiansen§ that in deriving the velocity constant of a bimolecular reaction involving ions, A and B say, the concentration terms have to be modified because of electrostatic effects which result in the concentration of B ions in the vicinity of an ion A being different from their bulk concentration. For *low concentrations* Christiansen has derived from kinetic considerations a suitable equation which, in addition, includes a steric factor that accounts for spatial orientation necessary for a fruitful collision, and a probability factor that allows for deactivation of the active complex formed from the reacting ions. For a reaction involving uni-univalent ions of opposite sign which form a complex X, the equation, exclusive of steric and probability factors is

$$\log k = \log \frac{Nr^2}{1000} \sqrt{8\pi RT \frac{M_A + M_B}{M_A M_B}} - \frac{E}{2.303RT} + \frac{2e^2\sqrt{2\pi N\mu}}{2.303(DKT)^{3/2}\sqrt{1000}},$$

N is Avogadro's number ( $6.06 \times 10^{23}$ ), M represents molar weight, R is the gas constant ( $8.315 \times 10^7$  ergs and 1.987 calories), T the absolute temperature,  $e$  the charge on a uni-univalent ion ( $4.77 \times 10^{-10}$  e.s.u.), D the dielectric constant of the solvent,  $k$  Boltzmann's constant ( $1.372 \times 10^{-16}$  ergs.),  $\mu$  the ionic strength, E the energy of activation, and  $r$  the distance between the centres of two colliding molecules that react. Since the last term in the

\* 'J. Chem. Soc.,' vol. 105, p. 690 (1914).

† Banks, Righellato, and Davies, 'Trans. Faraday Soc.,' vol. 27, p. 621 (1931).

‡ Davies, "The Conductivity of Solutions," p. 118 (1933).

§ 'Z. phys. Chem.,' vol. 113, p. 35 (1924).

equation is simply  $\log f_A f_B / f_X$ , and the other three terms are constant for a given solvent at a given temperature,  $\log k = \text{a constant} + \log f_A f_B / f_X$ , which is Brönsted's equation.

In order to apply the equation to methylammonium cyanate  $r$  may be assumed to be the sum of the ionic radii, calculated with the aid of the Stokes-Einstein formula,\* which connects the radius  $\rho$  of an ion with its mobility  $u$ :

$$\rho = k f / 6 N \pi \eta u.$$

$k$  the force acting on a gram ion is  $9650 \times 10^8$  dynes per centimetre,  $f$  the Faraday constant is 96500 Coulombs, and  $\eta$  the viscosity of water is given in the "International Critical Tables," vol. 5, p. 10. The sum of the mobilities of  $\text{CH}_3\text{NH}_3^+$  and  $\text{CNO}^-$  at  $40^\circ \text{C}$ . is about 161 (p. 295), and that of the mobilities of  $\text{K}^+$  and  $\text{CNO}^-$  at  $32^\circ \text{C}$ . is 152.† If the mobility of  $\text{K}^+$  at  $32^\circ \text{C}$ . is 84 and the temperature coefficient for  $\text{CNO}^-$  at  $18^\circ \text{C}$ . is 0.023, it follows that the mobility of  $\text{CNO}^-$  at  $40^\circ \text{C}$ . is 78 and that of  $\text{CH}_3\text{NH}_3^+$  83. The radii of the respective ions are therefore  $1.50$  and  $1.60 \times 10^{-8}$  cm., and  $r$  is  $3.10 \times 10^{-8}$  cm.  $E$  may be determined from Christiansen's equation by making appropriate substitutions for the temperatures  $39.9^\circ$  and  $50.0^\circ \text{C}$ . If  $r$  is assumed constant,  $D$  is taken to be the average of the accurate determinations of Drake, Pierce, and Dow,‡ and of Wyman,§ and the values of  $k$  corresponding to the molar concentration 0.0822 are employed, then  $E$  is 21,990 calories, which is 580 calories less than the Arrhenius value. Finally, the velocity constant calculated for infinite dilution, under which condition  $\log f_A f_B / f_X = 0$ , is  $k_0 = 5.0 \times 10^{-4}$  moles per litre per second. The experimental value in the same units is  $4.6 \times 10^{-4}$ , thus showing excellent agreement, despite the fact that steric effects and the possibility of deactivating collisions have been ignored.

If Christiansen's equation is a true representation of the mode of formation of methylurea from methylammonium cyanate one ought to be able to account similarly for the behaviour of other cyanates in water, and for the effect of change of solvent. Except for ammonium cyanate in aqueous solution the temperature coefficients of the reactions are unknown and velocity constants cannot be accurately evaluated. Of the alkylammonium cyanates investigated

\* 'Z. Elektrochem.,' vol. 14, p. 234 (1908).

† Walker and Kay, *loc. cit.*

‡ 'Phys. Rev.,' vol. 35, p. 613 (1930).

§ 'Phys. Rev.,' vol. 35, p. 623 (1930).

the rate of the fastest transformation, dimethylammonium cyanate, is three times that of the slowest, ethylammonium cyanate. As the molar weights of the cations are the same and the mobilities similar,\* the two factors involving  $r$  are not very different. A small difference in the value of  $E$ —700 calories—will account for the observed difference in rates. For all the cyanates examined by Walker and Appleyard (*loc. cit.*), a variation in  $E$  of 1000 calories, coupled with changes in molar weight and in  $r$ , will probably account for the experimental rates.

Change of solvent exerts a more pronounced effect, as is evinced by the transformation at 32° C. of ammonium cyanate (p. 302). For aqueous solution the critical increment (Arrhenius) is 23,170† calories, but the true critical increment may be less by 580 calories. The radius of  $\text{NH}_4^+$  in water is  $1.23 \times 10^{-8}$  cm. ( $u = 74$  at 25° C.), and the value of  $r$ , therefore,  $2.83 \times 10^{-8}$  cm. Hence at infinite dilution at 32° C. the velocity constant is  $8.9 \times 10^{-5}$  moles per litre per second, the experimental approximate value being  $9.9 \times 10^{-5}$ . For alcohol the Stokes-Einstein formula leads to a higher value of  $r$ . The mobility of  $\text{NH}_4^+$  at 25° C. is 19.2, whence ( $\eta = 0.01078$ ),  $\rho = 3.94 \times 10^{-8}$  cm. If the radius of  $\text{CNO}^-$  is supposed to be increased to the same extent as that of  $\text{CNS}^-$ , which has mobilities 65.4 and 29.2 in water and alcohol respectively at 25° C., its value is  $2.97 \times 10^{-8}$  cm., and  $r = 6.9 \times 10^{-8}$  cm. If  $E$  is taken to be the same as in water, substitution in Christiansen's equation ( $D = 23.3\dagger$ ) gives  $k_0 = 1.24 \times 10^{-3}$ , as compared with Ross's approximate value  $2.6 \times 10^{-3}$ . The agreement is not so good. It may be of significance that the use here of the value of  $r$  in aqueous solution leads to  $k_0 = 2.7 \times 10^{-3}$ , which is in excellent agreement. Otherwise a change in  $E$  of about 2000 calories will bring the rates of reaction in water and alcohol into line. Probably the radii calculated from the Stokes-Einstein formula are unsuitable for the reaction velocity equation. No doubt they refer to ions associated with sheaths of solvent molecules, whereas for reaction to occur the solvent layers must be broken down in order to allow the ions to approach to within a distance characteristic of solvent-free ions. This distance might well be independent of the solvent. As the values of  $r$  used above for aqueous solutions are comparable with the interionic distances measured in crystals of several uni-univalent salts they may be but slightly affected by ion solvation. The effect of solvation in increasing  $M_A$  and  $M_B$  and diminishing the

\* Lorenz and Posen, 'Z. anorg. Chem.,' vol. 94, p. 265 (1915).

† Moelwyn-Hughes, "The Kinetics of Reactions in Solution," p. 36 (1933).

‡ Åkerlöf, 'J. Amer. Chem. Soc.,' vol. 54, p. 4125 (1932).

collision frequency has been disregarded in the preceding calculations ; doubling  $M_A$  and  $M_B$  reduces  $k_0$  by only 30%.

It will be noted that constancy of  $r$ , coupled with constancy of  $E$ , involves the deduction that change of dielectric constant is mainly responsible for change in velocity in different solvents at high dilutions. If the collision frequency is not markedly affected by change of solvent the relationship between  $\log k_0$  and the reciprocal of the dielectric constant should be approximately linear. To a certain extent Walker and Kay's and Ross's results for ammonium cyanate in aqueous ethyl alcoholic solutions bear out this conclusion.

### *Conclusion.*

The reaction involving the transformation in aqueous solution of methylammonium cyanate into methylurea appears to be due to the interaction of methylammonium and cyanate ions, into which the salt is probably almost completely dissociated. The effect of changing ionic strength, whether due to the reactants or to added non-reacting electrolytes, is in keeping with this conclusion. As judged by the temperature coefficient the reaction is not a complicated one. Christiansen's equation for a reaction of measurable velocity between ions is applicable if it is assumed that every collision that leads to the formation of an active complex results in reaction. There is no obvious necessity to introduce a steric factor. The same mechanism probably accounts for the transformation of other substituted ammonium cyanates if the energies of activation vary a little (1000 calories) in the different reactions. In ethyl alcoholic solutions of ammonium cyanate, if the energy of activation is assumed to be the same as in water, a similar interpretation is possible, provided that the distance between the centres of ions that react to form a complex is about the same as in water. This distance may be independent of the solvent and similar to the nearest distance of approach of average unsolvated ions, or of uni-univalent ions in a crystal lattice. The dielectric constant of the solvent then plays an important part in determining the relative rates of reaction in different solvents. It must be emphasized, however, that there is no justification for assuming  $E$  to be independent of the solvent, and that accurate determinations of the energies of activation involved in these reactions in different solvents are required before definite conclusions can be drawn regarding the solvent effect. Experiments bearing on this point are contemplated.

It gives me great pleasure to accord my thanks to Sir James Walker, F.R.S., who proposed this research, and who has taken a keen interest in its develop-

ment. I am much indebted to the Trustees of the Carnegie Fund for the award of a Teaching Fellowship, and to the Trustees of the Moray Fund for several grants with which to purchase apparatus.

### *Summary.*

The rate of transformation of methylammonium cyanate into methylurea was determined in aqueous solution ( $< 0.1$  molar) at  $40^{\circ}$  and  $50^{\circ}$  C., and the temperature coefficient of the reaction was found. An extrapolation method had to be resorted to in order to determine the velocity constant at infinite dilution at  $40^{\circ}$  C., as a secondary reaction became prominent at low concentrations. Measurements of the equivalent conductivity of the solutions were made for the purpose of finding if the degree of dissociation, as calculated by Arrhenius from the conductivity ratio, entered into the reaction rate equation. It was found, however, that the rate conformed rather to an equation of the Brönsted type, involving thermodynamic activities. The effect of varying the concentration of the reactants and of adding electrolytes with and without common ions favoured this view.

The velocity constant calculated for an infinitely dilute solution at  $40^{\circ}$  C., with the aid of an approximate form of Christiansen's equation, for a bimolecular reaction of measurable velocity between ions, was in excellent agreement with the experimental result. Extension of the theory to the rates of transformation of other alkylammonium cyanates in water, and of ammonium cyanate in water and alcohol, led to the conclusion that the ions of the various cyanates were most probably directly responsible for the formation of the corresponding ureas, and that the dielectric constant of the solvent played a conspicuous part in determining the rate of reaction.

---

*The Reaction of Methane and Oxygen Sensitized by Nitrogen Peroxide. Part I—Thermal Ignition.*

By R. G. W. NORRISH and J. WALLACE, Department of Physical Chemistry  
Cambridge.

(Communicated by T. M. Lowry, F R.S.—Received January 20, 1934.)

Nitrogen peroxide has been shown to be effective in lowering the ignition temperature of several gases in air. Dixon\* found that 1% nitrogen peroxide gave a maximum lowering of the ignition temperature of methane in air, whilst only a trace of the peroxide lowered the ignition temperature of hydrogen in oxygen by as much as 200° C. This effect in the hydrogen-oxygen reaction has been studied by Dixon, Hinshelwood, and their co-workers, and by Norrish and Griffiths, and in the carbon monoxide-oxygen reaction by Semenov.†

Gibson and Hinshelwood‡ and Thompson and Hinshelwood§ working with hydrogen-oxygen mixtures found that over a wide range of temperature there are two pressures of nitrogen peroxide between which there is an explosion and above and below which there is a very slow reaction. Thus for small increasing pressures of nitrogen peroxide the reaction is catalysed, but beyond a certain concentration the nitrogen peroxide acts as an anti-catalyst; consequently the graph of ignition temperature against pressure of nitrogen peroxide showed a minimum. The catalytic effect was explained by Thompson and Hinshelwood as due to the initiation of chains by reaction between molecules of hydrogen and nitrogen peroxide to give ultimately activated molecules of hydrogen peroxide and branching chains leading to explosion; the inhibition was explained as being due to the breaking of the chains by the deactivation of the hydrogen peroxide molecules by collision with nitrogen peroxide molecules.

Norrish and Griffiths,|| however, investigating the hydrogen-oxygen reaction photosensitized and thermally sensitized by nitrogen peroxide proved that the chains are started by oxygen atoms and that they are broken by collision with the walls or with nitrogen peroxide. These oxygen atoms arise as a result of the direct dissociation of the nitrogen peroxide in the photochemical

\* 'Trans. Inst. Min. Eng.' vol. 80, p. 21 (1930-31).

† 'Z. phys. Chem.,' B, vol. 6, p. 307 (1930).

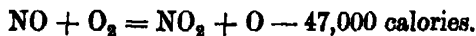
‡ 'Trans. Faraday Soc.,' vol. 24, p. 559 (1928).

§ 'Proc. Roy. Soc.,' A, vol. 124, p. 219 (1929).

|| 'Proc. Roy. Soc.,' A, vol. 139, p. 147 (1933).



reaction, while thermally it was suggested that they have their origin in the reaction



In the present work the effect of nitrogen peroxide on the ignition temperature of methane-oxygen mixtures has been studied with a view to obtaining more information as to the part played by the nitrogen peroxide. A mechanism is suggested, based on the initiation of chains by oxygen atoms arising from the above reaction.

### *Experimental.*

(a) *Furnace and Thermocouple.*—The electric furnace was made by winding evenly No. 22 gauge nichrome wire on a grooved quartz tube (internal diameter 5.3 cm.), four tappings being made at equal distances along the length of the furnace. The wire was covered with alundum cement and dried. Asbestos rope and finally asbestos board were wrapped round the tube, the tappings being protected by porcelain tubes. The ends of the furnace were closed by asbestos plugs. Through one of the latter passed the entry tube to the cylindrical quartz reaction vessel and the quartz sheath of the thermocouple, whilst through the other plug a small hole was bored and covered with a concave quartz lens. The temperature of the furnace was regulated by varying an external hand resistance.

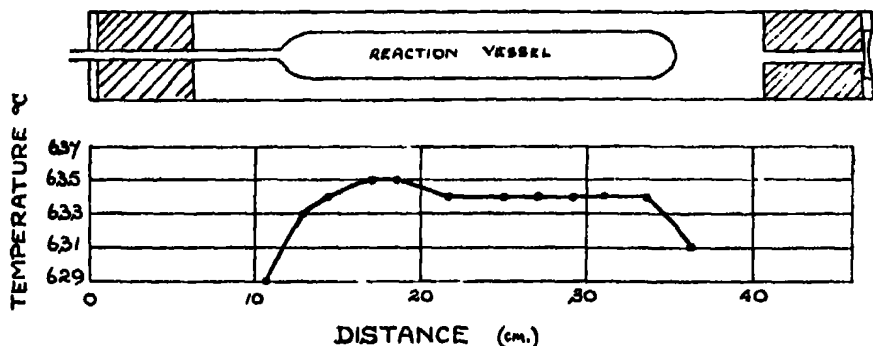


FIG. 1.—Uniformity of the temperature of the furnace.

Variable resistances were connected across the tappings of the furnace and so adjusted that under experimental conditions the furnace temperature was constant over as great a length of the tube as possible. Fig. 1 shows that at 634° C. the temperature was constant to within 1° C. over 20 cm.

The furnace temperature was measured by means of a chromel-alumel thermocouple. The couple had previously been calibrated using the melting points of three pure metals: antimony, 630° C.; zinc, 419° C.; lead, 327.5° C.

(b) *Reaction Vessel and Gas Supply*.—A sensitive Bourdon gauge B, fig. 2, connected to the reaction vessel was used as a null instrument to measure the pressure of methane and oxygen in a mixture, air being admitted to the outside of the gauge to balance the pressure. It was also used to measure small pressures of added nitrogen peroxide by reading the calibrated eyepiece scale of a microscope. The gauge was thermostated with water at room temperature. The bulb A was used as a mixing vessel, whilst in another bulb C a mixture of nitrogen peroxide and oxygen was made and could be suitably diluted when very small quantities of nitrogen peroxide, which were immeasurable on the gauge, were required in the final mixture in A. The mixing vessel A, the

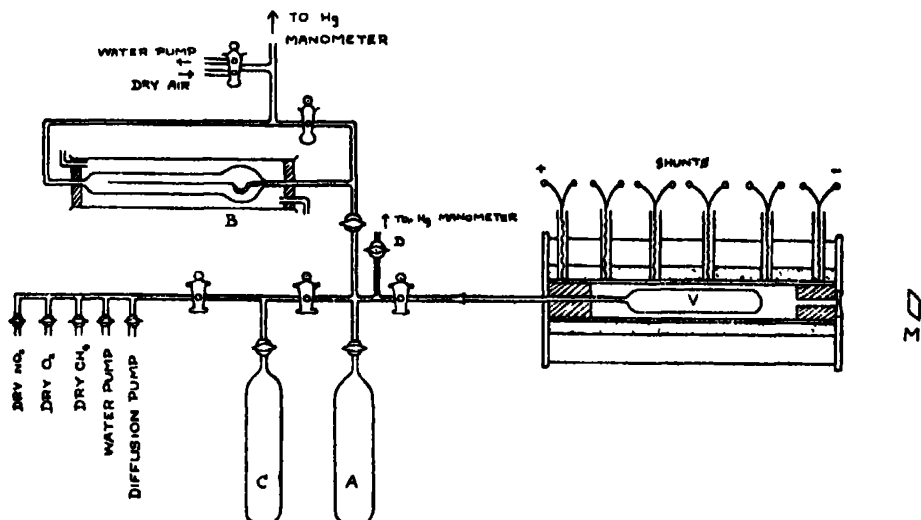


FIG. 2.

connecting tubing and the reaction vessel V were calibrated by using the mercury manometer D, so that the pressure of the gases in A being known, that of the gases when admitted to V could be calculated.

The apparatus could be evacuated first by a water pump and finally by a mercury diffusion pump backed by a Toepler pump.

(c) *Gases*.—Methane was prepared by the action of a zinc-copper couple on methyl iodide in methyl alcohol at 60° C. The gas was washed by a solution of sodium methoxide in methyl alcohol to remove unchanged methyl iodide and afterwards by 20% fuming sulphuric acid to remove any unsaturated hydrocarbons. After collecting the gas over a dilute solution of sodium hydroxide to remove sulphuric acid fumes and any iodine (which was shown by Dixon to have a very marked effect in raising the ignition temperatures of

methane in oxygen) it was fractionated twice after liquefying with liquid air and finally collected in an aspirator over air-free water.

Oxygen was obtained from a cylinder and nitrogen peroxide from a stock kindly supplied by I.C.I., Ltd. The nitrogen peroxide was freed from traces of nitrogen trioxide by nearly solidifying the gas and pumping away several fractions until the resulting solid was colourless. All gases were dried by passing them over pure granulated phosphorus pentoxide.

(d) *Technique*.—To determine an ignition temperature the furnace temperature was regulated, the mixture prepared, and the reaction vessel evacuated. The tap of the mixing vessel was opened and the mixture allowed to enter the reaction vessel. At high total pressures the explosion was noted by a yellow flash in the connecting tubing, but at low pressures it was observed by looking into the inclined mirror M at the reflection of the end of the reaction vessel in the furnace. The ignition temperature was bracketed between two temperatures. At one there was an explosion and at the other no explosion. The range was then narrowed until at  $T^{\circ}\text{C}$ . there was an explosion and at  $T-5^{\circ}\text{C}$ . no explosion. The experiments at these temperatures were repeated and the mean temperature  $\frac{2T-5}{2}^{\circ}\text{C}$ . was taken as the ignition temperature of the mixture. The lag between the time of opening the tap and the appearance of the flash was recorded by means of a stop-watch.

It was found that if the first run of the day did not give an explosion it was not to be relied upon. This may have been due to the change in the conditions of the surface of the reaction vessel overnight.

On several occasions using high pressures of mixtures the four-way piece shown in the diagram was blown to pieces.

#### *The Effect of Variation of Pressure of Nitrogen Peroxide on the Ignition Temperature.*

Throughout the present work equimolecular mixtures of methane and oxygen were used. Since in comparing a series of experiments involving a changing explosion temperature it is important to keep a constant molecular concentration, care was taken so to adjust the pressure in accordance with the temperature at which it was proposed to work. This was done by calculating from the known relative volumes and temperatures of the mixing vessel and connections and the reaction vessel, the pressure needed to give the required final concentration in the reaction vessel.

Three series of experiments were carried out at concentrations corresponding to total pressures of 229.7 mm., 88.6 mm., and 360.0 mm. at 500° C. The ignition temperatures were determined as the pressure\* of nitrogen peroxide was increased from 0.0 mm. to 25.0 mm. The results are recorded in Tables I, II, and III and are represented graphically in fig. 3.

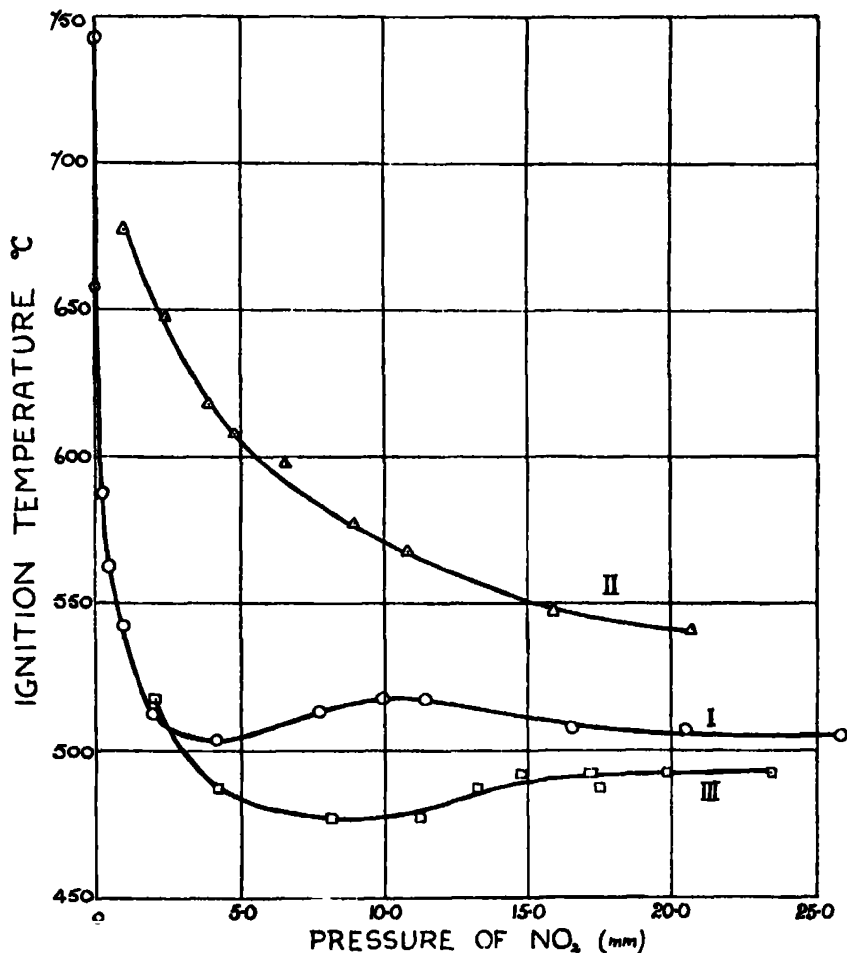


FIG. 3.

It can be seen that in the case of the concentration corresponding to 229.7 mm. at 500° C. the ignition temperature of the mixture is lowered 239.5° C. by 4.17 mm. of NO<sub>2</sub>. As the pressure of nitrogen peroxide increases the catalytic effect of the sensitizer becomes constant. It will also be noticed

\* The pressures recorded are the actual pressures in the reaction vessel at the temperature of ignition, representing all the added nitrogen peroxide as NO<sub>2</sub>.

Table I.—Summary of limits for constant concentration of methane and oxygen corresponding to total pressure  $\text{CH}_4 + \text{O}_2 = 229.7$  mm. at  $500^\circ \text{C}$ . (Runs 1-61.)

Limit.	Pressure* $\text{NO}_2$ .	Lag.	Ignition temperature.
	mm.	sec.	$^\circ \text{C}$ .
A	0.00	3.2	743
M	0.07	1.3	857.5
L	0.25	2.3	587.5
K	0.48	4.3	562.5
J	0.98	2.0	542.5
G	1.98	4.5	512.5
B	4.17	6.3	503.5
D	7.72	3.4	513.5
I	9.91	1.9	517.5
C	11.43	1.3	517.5
E	16.52	1.3	507.5
F	20.48	0.0	507.5
H	25.82	1.5	505.5

\* See footnote p. 311.

Table II.—Summary of limits for constant concentration of methane and oxygen corresponding to total pressure  $\text{CH}_4 + \text{O}_2 = 88.6$  mm. at  $500^\circ \text{C}$ . (Runs 1A to 88A.)

Limit.	Pressure* $\text{NO}_2$ .	Lag.	Ignition temperature.
	mm.	sec.	$^\circ \text{C}$ .
F'	1.04	0.0	877.5
E'	2.49	0.0	847.5
J'	3.92	0.0	817.5
I'	4.80	0.0	807.5
H'	6.59	0.0	597.5
D'	7.27	0.0	582.5
G'	8.96	0.0	577.5
A'	10.83	1.1	567.5
C'	15.92	1.1	547.5
B'	20.74	0.0	541.5

\* See footnote p. 311.

that with high total concentrations of methane and oxygen the lowering of ignition temperatures by nitrogen peroxide is more effective than with low total concentrations of the combustible gases.

*The Effect of Surface on Ignition Temperature.\**—Five cylindrical quartz reaction vessels of the same length but of different diameters, viz., 3.44, 2.53,

\* We are indebted to Mr. E. I. Akeroyd of this laboratory for valuable assistance in completing this series of experiments.

Table III.—Summary of limits for constant concentrations of methane and oxygen corresponding to total pressure  $\text{CH}_4 + \text{O}_2 = 360$  mm. at  $500^\circ \text{C}$ . (Runs 1B to 42B.)

Limit.	Pressure* $\text{NO}_2$ .	Lag.	Ignition temperature.
	mm.	sec.	$^\circ \text{C}$ .
J''	2.06	1.9	517.5
A''	4.23	5.5	487.5
F''	8.13	9.3	477.5
B''	11.26	9.0	477.5
I''	13.24	8.5	487.5
E''	14.78	6.9	492.5
H''	17.22	1.5	492.5
C''	17.50	6.7	497.5
G''	19.82	2.9	492.5
D''	23.45	0 0	492.5

\* See footnote p. 311.

2.28, 2.00, and 1.52 cm., were used in these experiments. The results are given in Table IV. Two series of experiments were carried out: series A at high concentration (corresponding to a total pressure of methane and oxygen equal to 229.7 mm. at  $500^\circ \text{C}$ .); series B at low concentration (corresponding to a total pressure of 88.6 mm. at  $500^\circ \text{C}$ .). The following results were obtained:—

Table IV.

Diameter of reaction vessel.	Ignition temperature.	
	Series A.	Series B.
	$p(\text{NO}_2) = 4.10$ mm. $^\circ \text{C}$ .	$p(\text{NO}_2) = 8.80$ mm. $^\circ \text{C}$ .
3.44	487.5	552.5
2.53	503.5	597.5
2.28	505	657.5
2.00	515	752.5
1.52	537.5	Greater than $940^\circ$

These results are plotted in fig. 4. It will be seen that the ignition temperature is dependent upon the diameter of the vessel in a very remarkable manner. At the higher pressure (Series A) there is a small but progressive increase of the ignition temperature with decreasing diameter; at the lower pressure (Series B) the ignition temperature after first slowly increasing with decrease of diameter becomes remarkably sensitive to further change, increasing very rapidly with further decrease of diameter.

The interpretation of these extraordinary facts can only be satisfactorily achieved on the basis of a chain mechanism. As we shall see, several factors may influence the length of chains and consequently the ignition temperature, one of which is the diffusion of reaction centres to the surface. These factors may all operate simultaneously, but under certain experimental conditions (reduced pressure) the surface factor may become predominant. This corresponds to the marked effect of diameter noted in Series B above.

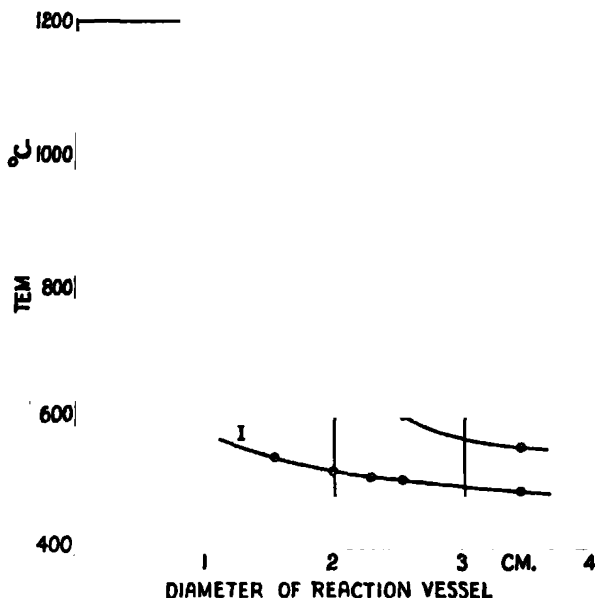
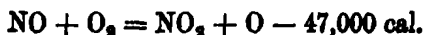


FIG. 4.

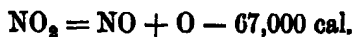
### Discussion.

*Mechanism of the Reaction.*—It has been shown\* that reaction chains in the hydrogen oxygen reaction are generated by oxygen atoms directly formed from nitrogen peroxide by photochemical decomposition. It has been further demonstrated that the kinetics of the reaction thermally sensitized by nitrogen peroxide are identical with the kinetics of the photochemical reaction, and there is little doubt that in this latter case also, the function of the sensitizer is to provide the necessary atoms of oxygen to start the chains. The main point of uncertainty is the mechanism by which these oxygen atoms are generated thermally. Norrish and Griffiths considered that their most probable source was in the reaction



\* Norrish and Griffiths, 'Proc. Roy. Soc.,' A, vol. 139, p. 147 (1933).

but the reaction



suggested by Schumacher\* though less probable in view of the greater energy absorption involved is not excluded.

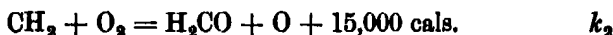
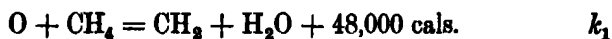
It is not yet possible to distinguish definitely between these two possibilities, but the facts of this paper are well represented if we write

$$\frac{d[\text{O}]}{dt} = k[\text{NO}_2].$$

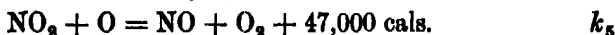
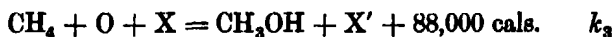
As a provisional assumption we shall assume that reaction is propagated from the oxygen atoms by a simple chain mechanism involving  $\text{CH}_2$  radicals.

This assumption of the presence of  $\text{CH}_2$  radicals is in accord with the experiments of Belchetz† which show that  $\text{CH}_2$  is more readily formed from methane than  $\text{CH}_3$ , and with theoretical considerations,‡ which indicate a maximum probability for the formation of the  $\text{CH}_2$  radical in a singlet state. The following scheme is assumed:—

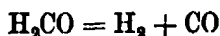
(a) *Chain propagation*—



(b) *Chain ending*—



According to this scheme, the combustion of the methane proceeds with the formation of water and formaldehyde. The latter is rapidly decomposed at the temperature of the explosion



so that the final stoichiometric equation is



as established by Bone and Drugman.§

\* 'Z. phys. Chem.,' B, vol. 10, p. 7 (1930).

† 'Trans. Faraday Soc.,' vol. 30, p. 89 (1934).

‡ Norrish, 'Trans. Faraday Soc.,' vol. 30, p. 103 (1934).

§ 'J. Chem. Soc.,' vol. 89, p. 660 (1906).



The chains are terminated by oxygen atoms (1) when they reach the surface of the vessel, (2) when a ternary collision leads to the formation of methyl-alcohol, or (3) when they are removed by reaction with nitrogen peroxide. With the exception of the reaction starting the chains, all the others involve atoms or radicals, and are highly exothermic. We shall therefore assume that they require but slight activation and are virtually independent of temperature in comparison with the endothermic reaction responsible for starting the chains.

It is then easy to deduce the following kinetic expression for the rate of oxidation of methane

$$-\frac{d[\text{CH}_4]}{dt} = \frac{kk_1 e^{-E/RT} [\text{NO}_2][\text{CH}_4]}{k_3 [\text{X}][\text{CH}_4] + k_5 [\text{NO}_2] + k_4 [\text{S}]},$$

where  $[\text{S}]$  represents the factor for surface deactivation, and  $k_3 [\text{X}]$ , the factor for ternary collisions, is small compared with  $k_1$  the factor for ternary collisions, of  $\text{CH}_4$  with O atoms.

When the reaction rate reaches a certain limiting value the heat produced by the highly exothermic processes of the chain cannot be removed rapidly enough and the process becomes self-heating. We take this as the condition which must be satisfied for inflammation. Thus for explosion,

$$-\frac{d\text{CH}_4}{dt} = \text{const.} = K.$$

It follows that

$$\frac{K (k_3 [\text{X}][\text{CH}_4] + k_5 [\text{NO}_2] + k_4 [\text{S}])}{kk_1 [\text{NO}_2][\text{CH}_4]} = e^{-E/RT}.$$

Thus,

$$T = \frac{E}{R \log \frac{kk_1 [\text{NO}_2][\text{CH}_4]}{K (k_3 [\text{X}][\text{CH}_4] + k_5 [\text{NO}_2] + k_4 [\text{S}])}}.$$

This gives the temperature of ignition as a function of concentrations of nitrogen peroxide and methane. For any one of the curves of fig. 3 the concentration of methane was constant, and the term  $k_4 [\text{S}]$  was also nearly constant since the same reaction vessel was used throughout. Thus with low concentrations of nitrogen peroxide, when  $k_5 (\text{NO}_2)$  is negligible in comparison with  $k_3 (\text{X})(\text{CH}_4)$  and  $k_4 (\text{S})$ ,

$$T \propto \frac{1}{\log [\text{NO}_2]}.$$

This relationship has been tested for the limits B, G, J, K, L, and M of Table I as shown in fig. 5, where it is seen to be sensibly true, a good straight line being obtained when  $1/T$  is plotted against  $\log [\text{NO}_2]$ .

With higher concentrations of nitrogen peroxide, the above relationship can no longer hold true, for  $k_5$   $[\text{NO}_2]$  is now no longer negligible, and indeed finally becomes the controlling factor. When this is so the ignition temperature becomes independent of concentration of nitrogen peroxide (curves I and III, fig. 3), and the general expression reduces to

$$T = \frac{E}{R \log \frac{k k_1}{k_5} [\text{CH}_4]}$$

Thus

$$T \log [\text{CH}_4] = \text{const.},$$

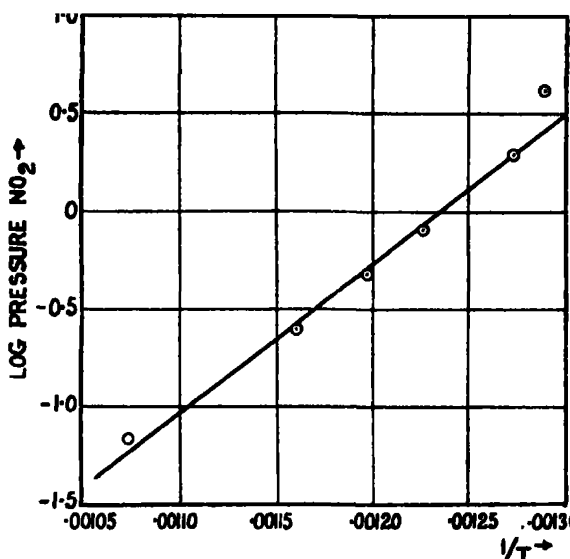


FIG. 5.

and the greater the concentration of methane the lower the explosion limit. That the above relationship is approximately true can be seen from the following table, in which the data are taken from Tables I, II, and III.

Table V.

Concentration of $\text{CH}_4$ mm. at 500° C.	Limiting temperature of explosion with high $[\text{NO}_2]$ .	$\log [\text{CH}_4] \times T$ .
mm.	° C.	
44	535	133
115	505	160
180	492	172

It will be seen that with curve II, fig. 3, which corresponds to low total concentration of gas, the limiting ignition temperature is more slowly reached. This is due to the fact that diffusion occurs more readily at low pressure, and the factor  $k_4 [S]$  therefore assumes a greater relative importance in comparison with  $k_5 [NO_2]$ ; as a consequence the explosion limit is now given by the expression

$$T = \frac{E}{R \log \frac{kk_1 [NO_2] [CH_4]}{K (k_5 [NO_2] + k_4 [S])}}$$

and does not become independent of  $[NO_2]$  until the latter reaches much higher concentrations than with curves I and III.

The importance of the surface factor is further well illustrated by curves I and II fig. 4. It will be seen that with high total pressures of gas, the ignition limit is nearly independent of the diameter and therefore of any influence of the surface of the reaction vessel (curve I). With low total pressures, on the other hand (curve II), the ignition temperature is very sensitive to the variation of diameter and is thus strongly dependent upon the surface factor. If surface deactivation were the sole controlling factor, we should have, for constant  $[NO_2]$  and  $[CH_4]$

$$T \propto \frac{1}{\log \frac{1}{[S]}}$$

The surface deactivation factor may be set proportional to (1) the diffusion coefficient ( $D$ ) of the reaction centres in the medium and (2) to the reciprocal of the diameter ( $d$ ) of the reaction vessel. Thus

$$[S] \propto \frac{D}{d}.$$

(Justification for this may be found in the fact that the rate of removal of active centres from any small element in the axis of the cylindrical reaction vessel is proportional to the product of the diffusion coefficient and the concentration gradient of active centres between the axis and the surface. When the surface is the sole controlling factor, this latter can be shown to be linear, and thus for the limiting concentration of centres in the axial element corresponding to explosion, will be inversely as the diameter of the vessel.)

When the concentration is kept constant as in the curves of fig. 4.

$$D \propto \sqrt{T},$$

where  $T$  is the absolute temperature.

Thus

$$[S] \propto \frac{\sqrt{T}}{d},$$

and we have for the ignition temperature

$$T \log \frac{1}{[S]} = T \log \frac{d}{\sqrt{T}} = \text{const.}$$

That this is approximately true is seen from Table VI which is derived from the experimental data of curve (II) fig. 4.

Table VI.

Diameter of vessel.	Explosion temperature ° abs.	$T \log d/\sqrt{T}$ .
3.45	825.5	89
2.52	870.5	81
2.28	930.5	81.5
2.00	1025.5	81.5
1.52	> 1213	—
1.52	Extrap. 1398	(85)

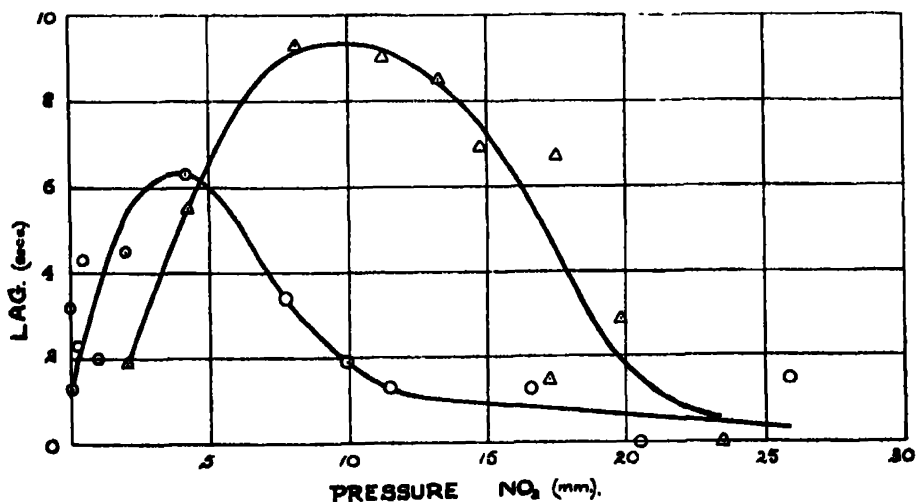


FIG. 6.

*Induction Periods.*—The time lags before ignition of the gases recorded in Tables I-III were plotted against the pressures of nitrogen peroxide in the mixtures in fig. 6.

These lags are immeasurable by means of a stop-watch at low concentrations of the mixtures, but with the concentrations corresponding to pressures of 229.7 mm. and 360.0 mm. at 500° C. they could be measured. The curves show maxima which correspond with the minima in curves I and III of fig. 3.

According to the kinetic mechanism outlined above, explosion occurs when the velocity of methane decomposition reaches a limiting magnitude. This will be dependent on the attainment of a definite partial pressure of oxygen atoms. At high concentrations of nitrogen peroxide, when the explosion temperature is constant, the time required for this condition to be realized will be greater the smaller the concentration of nitrogen peroxide. As the concentration is further reduced, however, the temperature of explosion rises rapidly; this will increase the rate of generation of oxygen atoms, with the result that the time required to generate the required number decreases. Thus the induction periods pass through a maximum.

*Conclusion.*—This study of the kinetics of the sensitized explosion of methane and oxygen shows that the kinetics can be adequately explained on the basis of a simple chain mechanism. While the exact nature of this chain is yet to be established, we have shown that all the facts at present available are explainable by the very simple scheme suggested. This applies not only to the kinetic studies but also to the analytical results of Bone and his co-workers. Further direct evidence for the participation of  $\text{CH}_3$  radicals in the chain is, however, necessary before we can be sure of the mechanism, and we shall continue to seek carefully for such evidence. If it should happen that later evidence indicates a different chain carrier, any change in the mechanism must be in accord with the general kinetics established above—i.e., chains start from oxygen atoms generated from nitrogen peroxide, and finish either by (1) adsorption by the surface, (2) reaction with nitrogen peroxide, or (3) by some process involving a ternary collision.

We are indebted to the Royal Society, The Chemical Society, and to Imperial Chemical Industries, Ltd., for grants, and to the Department of Scientific and Industrial Research for a maintenance grant for one of us (J. W.).

(1) The ignition temperatures of equimolecular constant concentrations of methane and oxygen with varying concentrations of nitrogen peroxide have been determined by a static method. Three constant concentrations of combustible gases were used.

(2) The effect of surface on the reaction has been studied.

(3) The length of the induction period has been observed to pass through a maximum with increasing pressure of nitrogen peroxide.

(4) A reaction mechanism has been suggested. Oxygen atoms arising from the nitrogen peroxide are assumed to start chains in which  $\text{CH}_3$  radicals and oxygen atoms are concerned. These ultimately lead to explosion of the gases when a limiting velocity involving self-heating of the mixture is reached. The theory is in agreement both with the kinetic results of the present paper and with the analytical data of Bone and others relative to the products of explosion and slow reaction.

---

### *Band Spectrum of Aluminium Chloride ( $\text{AlCl}_3$ ).*

By B. N. BHADURI, M.Sc., Research Student, Imperial College, South Kensington, and A. FOWLER, F.R.S., Yarrow Research Professor of the Royal Society.

(Received February 17, 1934.)

[PLATE 8.]

#### *Introductory.*

In the course of experiments on the spectrum of hydrogen chloride, a strong band system having its greatest intensity about  $\lambda$  2610 was observed when using discharge tubes with aluminium electrodes. It was found that these bands had previously been noted by Jevons\* as occurring in his experiments on silicon chloride in active nitrogen, and had subsequently been described and attributed by him† to aluminium chloride. The photographs obtained by the authors, however, showed this band system in far greater detail than those of Jevons, and as it obviously included a large number of bands associated with the heavier isotope of chlorine, it seemed desirable to investigate the structure of the spectrum more completely.

For the more complete production of this band system, Jevons passed discharges through the vapour of aluminium chloride,  $\text{AlCl}_3$ , and this method has been adopted in the present investigation. As obtained in this way, the

\* 'Proc. Roy. Soc.,' A, vol. 89, p. 187 (1913).

† 'Proc. Roy. Soc.,' A, vol. 106, p. 186 (1924).

band system extends from  $\lambda$  2555 to about  $\lambda$  2810 and exhibits features of considerable interest besides those due to the presence of isotopes.

The experiments made by Jevons left no doubt that the bands under consideration originated in molecules containing aluminium and chlorine. Thus, the bands were not detected in the spectra of discharges between aluminium electrodes in gases other than chlorine, nor in chlorine between electrodes of gold. The vibrational analysis included in the present paper further demonstrates conclusively that the bands are due to diatomic molecules  $\text{AlCl}$ .

### *Experimental.*

The discharge tube employed was of ordinary H form, having a capillary about 1 mm. in diameter; it was made of pyrex glass and was provided with a quartz window to permit observations in the ultra-violet. The electrodes were of aluminium. A side tube contained the (solid) aluminium chloride  $\text{AlCl}_3$ , and a liquid air trap and a tube containing caustic soda were introduced to prevent chlorine reaching the pump.

After the exhaustion of air, the discharge tube was readily filled with the vapour  $\text{AlCl}_3$  by gently heating the container with a Bunsen flame. On the displacement of air by the vapour, the colour of the discharge became pale blue, and there was no difficulty in maintaining a fairly constant pressure by continuous pumping and appropriate heating of the chloride. The vapour was pumped away from the window end of the discharge tube so as to minimize the solid deposit on the window itself.

The discharge was supplied by a Ferranti transformer, the primary being at 80 volts and taking a current of 3 amperes. An uncondensed discharge was used throughout. The discharge tube was thoroughly "cleaned" by pumping and heating before taking photographs, and the spectra were remarkably free from lines or bands due to impurities.

After a few preliminary photographs had been taken with smaller instruments, Hilger's E1 quartz spectrograph was used exclusively. With this instrument the dispersion ranges from 4 Å. per millimetre at  $\lambda$  2810 to 3 Å. per millimetre at  $\lambda$  2600, and the definition and resolution were very satisfactory. Exposures varied from 5 minutes for the brightest region about  $\lambda$  2610 up to 3 hours for the fainter bands on both sides of this. Ilford ordinary plates were used.

The iron arc was used for reference spectra, standard wave-lengths being taken, so far as possible, from the tables published by the International Astronomical Union.\* When these were inadequate in number for satisfactory

\* 'Trans. Int. Astr. Union,' vol. 4 (1933).

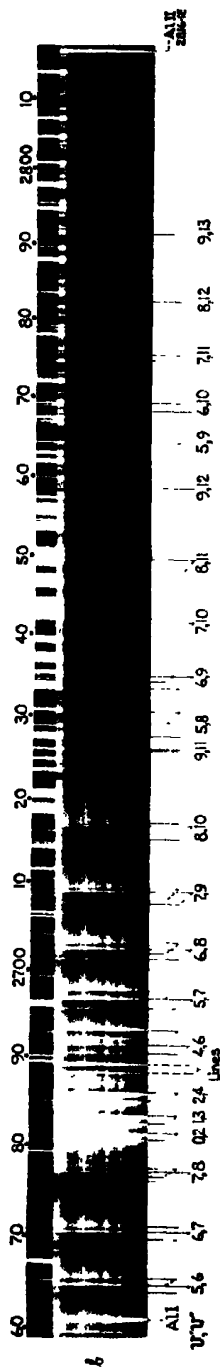




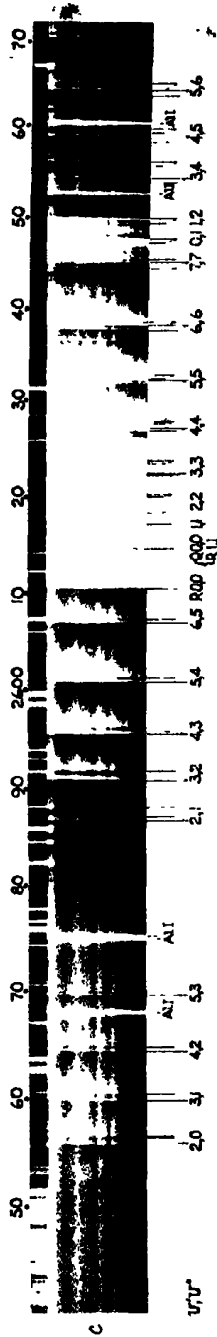
GENERAL APPEARANCE OF  $\text{AlCl}$  BANDS



BANDS AT THE LESS REFRACTIBLE END



BANDS AT THE MORE REFRACTIBLE END



a.—General appearance of  $\text{AlCl}$  bands,  $\lambda$  2810– $\gamma$  2555. b.— $\text{AlCl}$  bands towards the less refrangible side,  $\lambda$  2810– $\gamma$  2660. c.— $\text{AlCl}$  bands towards the more refrangible side,  $\lambda$  2670– $\gamma$  2555. Long and short vertical lines represent R and Q heads respectively, the associated heads due to the less abundant isotope are represented by shorter lines and are connected together by dotted lines.

interpolations, they were supplemented by interferometer values for the vacuum iron arc given by Burns and Walters.\*

Although the dispersion and resolution on the photographs may have appeared sufficient to justify an attempt to determine wave-lengths approximately correct to the third decimal place, it has not been considered desirable to express the wave-lengths to more than two decimal places. Experience has shown, in fact, that when the plate is at a large angle to the spectrum, as in a quartz spectrograph, small errors are liable to be introduced by the lack of perfect flatness of the plates. Further, there is usually some uncertainty as to correct settings on the heads of bands, and different observers might get slightly differing wave-lengths. The results of the vibrational analysis, however, indicate that the wave-lengths are mostly correct to within two or three hundredths of an angstrom.

### *Description of the Spectrum.*

The spectrum cannot be adequately described without reference to photographs, some of which are reproduced in Plate 8. Spectrum *a* gives a general view of the whole system of bands, with the exception of a few faint bands towards the red limit, which have been omitted in order to permit the representation of the spectrum on a somewhat larger scale. Spectra *b* and *c* show different regions on a still larger scale.

It will be seen that the great majority of the bands are shaded towards the red and are spaced with obvious regularity. There are, however, two strong short sequences of bands which are shaded in the opposite direction, the least refrangible heads being at  $\lambda$  2685.68 and  $\lambda$  2649.73 respectively. The heads of most of the bands are very sharply defined, but the rotational structure is too fine for resolution except in a few bands, and especially in those which appear on the more refrangible side of the strong group at  $\lambda$  2610, where the dispersion is greatest.

The principal band heads occur in pairs and are due to the more abundant molecules  $\text{AlCl}^{35}$ . These, when bright enough, are duplicated by fainter pairs which represent the less abundant molecules  $\text{AlCl}^{37}$ . As will appear later, the origin of the band system is near  $\lambda$  2613.4, on the red side of the great band head at  $\lambda$  2610, and the heavier isotope bands are displaced from the brighter bands due to  $\text{AlCl}^{35}$  by amounts which are roughly proportional to the respective distances of the various bands from the origin of the system.

\* 'Pub. Allegheny Obs.,' vol. 8, No. 4 (1931).

Thus, on the red side of the origin, the displacements are towards the extreme ultra-violet and range from 32 units of wave number at  $\lambda$  2810 to less than one unit at the  $\lambda$  2617 band, which is near the system origin. On the other hand, on the more refrangible side of the origin, the bands due to the heavier molecules are displaced towards the red, the displacement amounting to 9 wave number units at  $\lambda$  2555 where the last band is observed. These features will readily be noted on examination of the photographs, and they are indicated quantitatively in the tables which follow.

The general appearance of the double-headed bands at once suggests that the (usually) stronger components represent R branches, whilst the fainter ones, which have equally sharp heads, represent Q branches. P branches, which do not ordinarily form sharp heads (in bands shaded to the red) may be superposed on the Q branches, but this cannot be determined for want of adequate resolution of the rotational structure. For the purposes of the present paper it will suffice to designate the probably combined P and Q branches as Q. In general the R bands extend only a short distance from their heads and show no signs of structure with the dispersion employed, whilst the Q bands extend further and are in places partially resolved.

The appearance of the spectrum in the main group is considerably modified in some of the photographs on account of absorption by the luminous vapour which occupied the space between the end of the capillary of the discharge tube and the quartz window through which the spectrum was observed. Some of the band heads were definitely reversed. In the present paper, however, an attempt has been made to eliminate these absorption effects, so that the data presented refer exclusively to the spectrum of emission bands. The peculiarities of the absorption effects will form the subject of a further investigation.

Another modification of the main group of bands was found when the intensity of the discharge was considerably reduced. In particular, the head at  $\lambda$  2610 then became very feeble, as in one of the photographs reproduced by Jevons. This is the R 0, 0 band and its head is at a considerable distance from the band origin; it may therefore be supposed that with low excitation the rotation band lines do not extend sufficiently far from the band origin to reach the head with any great intensity. In bands further removed from the system origin the heads are nearer to their respective band origins and appear fully developed.

In addition to the system of bands which has been described, Jevons observed eight bands, also degraded towards the red, in the region  $\lambda$  3596 to  $\lambda$  3111. The heads were widely separated and were of sufficiently open structure to



To save p. 325).



be partially resolved even with very small dispersion. These bands, however, were not observed in the present experiments, and as they have no apparent relation to the  $\text{AlCl}$  bands, they may possibly have originated in some impurity.

### *Vibrational Analysis.*

Most of the bands were easily arranged in a  $v' v''$  scheme representing transitions between initial vibrational levels  $v'$  and final levels  $v''$ . Consideration of intensities suggested that the most refrangible head in the main group at  $\lambda 2610.23$  was to be taken as R 0, 0, whilst a band at  $\lambda 2614.44$  was Q 0, 0. This assignment of the vibrational quantum numbers was subsequently confirmed by the close agreement between the calculated and observed magnitudes of the isotope displacements when the origin of the system had been determined.

The distribution of the band heads is exhibited in the usual manner in Table I. The bands are there arranged in groups such that a complete group includes the main R and Q heads and the associated isotope heads  $\text{R}'$  and  $\text{Q}'$ , in the order Q,  $\text{Q}'$ , R and  $\text{R}'$ . Some of the groups are incomplete on account of the low intensity of one or more members, and others because of their proximity to the system origin and the consequent very small displacements of the isotope heads.

The differences  $\nu_{v'} - \nu_{v'-1}$ , and  $\nu_{v''} - \nu_{v''-1}$ , are shown in Table I for both R and Q heads. It will be seen that the differences for the Q heads are very uniform in the same vertical column or horizontal row, proving that the measured positions of the Q heads are very close to the respective band origins. The corresponding differences for the R heads, as was to be expected, show no such uniformity, on account of the varying distances of the R heads from their band origins; the variations, however, are quite systematic.

Assuming then that the Q heads represent band origins, the origin of the system can be calculated in the usual way by adjusting the observed wave numbers of these heads to the well-known general formula\*

$$\begin{aligned} \nu_0 = \nu_s + [\omega'_s (v' + \tfrac{1}{2}) - x'_s \omega'_s (v' + \tfrac{1}{2})^2 + y'_s \omega'_s (v' + \tfrac{1}{2})^3] \\ - [\omega''_s (v'' + \tfrac{1}{2}) - x''_s \omega''_s (v'' + \tfrac{1}{2})^2], \end{aligned}$$

where  $\nu_0$  represents band origins and  $\nu_s$  the origin of the system. A cubic term is quite unnecessary for the final levels ( $v''$ ) in  $\text{AlCl}$ , as clearly appears from the uniformity of the second differences between successive Q bands.

\* See Jevons' "Report on the Spectra of Diatomic Molecules," 'Phys. Soc.,' p. 56 (1932).

The mean value of the second differences is 3.9 units of wave number, and the half of this, 1.95, has been adopted for the coefficient  $x''$ ,  $\omega''$ . With this value, the bands 2, 0 and 2, 2 gave  $\omega''$ , as 481.20, while 9, 11 and 9, 13 gave 481.40; taking the mean of these, 481.30 was adopted for  $\omega''$ .

Next, by a graphical extrapolation of the main sequence (0, 0; 1, 1; etc.) the origin of the system,  $v_0$ , was found to be very near to  $v$  38254. With this and the constants  $\omega''$ , and  $x''$ ,  $\omega''$ , already determined, values of  $\omega'$ ,  $x'$ ,  $\omega'$ , and  $y'$ ,  $\omega'$ , were calculated from the three Q heads 0, 0; 5, 6; and 9, 12. The resulting equation for the Q heads of the whole system was thus found to be:

$$v_0 = 38254.0 + [449.96 (v' + \frac{1}{2}) - 4.37 (v' + \frac{1}{2})^2 - 0.216 (v' + \frac{1}{2})^3] \\ - [481.30 (v'' + \frac{1}{2}) - 1.95 (v'' + \frac{1}{2})^2].$$

With these values of the constants, the agreement between the calculated and observed wave numbers of the Q heads is as close as can be expected and the sums of the positive and negative residuals are so nearly equal as to justify the retention of 38254.0 for the system origin.

Table II.—Observed-Computed Wave numbers of Q Heads.

$v''$	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.0*	0.0	+0.3											
1	-0.6	-0.3	+1.2	+1.4										
2	-0.4*	-0.6	-0.3*		+1.7									
3		-0.1	-1.1	-0.3	-0.7									
4			-1.0	-1.0	-1.0	-0.3	+0.8							
5				-1.4	-1.1	-1.4	0.0*	+0.7	+0.9	-1.3	+3.7			
6					-0.9	-0.6		+0.6	+1.0	-1.6	+2.9	+3.0		
7								+0.5	+0.7	-1.1	+2.0	+2.2	+1.8	
8											+1.3	+1.8	+1.7	
9												-0.1*	-0.2*	-0.3*

\* Used in calculation of constants.

It is of some interest to note the extent to which such an equation is capable of representing the band origins in an extended system. The observed *minus* calculated values for the Q heads are accordingly collected in Table II, where the bands used in the calculation of constants are marked with an asterisk.

The deviations from the formula are due in part to errors of observation, but it will be noted they are mostly of a systematic character and thus indicate the general accuracy of the observational data. There is, however, a marked



discontinuity of the residuals at  $\nu'' = 9$  which suggests a small perturbation of this level. For the 47 heads the mean deviation, regardless of sign is  $+1.1$  while the algebraical mean is  $+0.1$ , and it is clear that the formula is sufficiently exact for the calculation of isotope displacements.

*Bands Degraded away from the Red.*

An unusual result of the vibrational analysis is that the strong bands degraded away from the red, which have previously been mentioned, find places in the  $\nu' \nu''$  scheme for Q bands which are shaded towards the red. These also are accompanied by isotope bands which are displaced from the main bands by amounts in close agreement with those calculated on the supposition that they belong to the system having its origin at  $\nu 38254$ . In Table III the observational data are compared with the values calculated from the general formula and from the formula for isotope displacements which is given later.

Table III.—Bands Degraded away from the Red.

$\nu' \nu''$ .	$\lambda$ (Int.).	$\nu$ obs.	$\nu$ calc.	Isotope displacement.	
				Observed.	Calculated.
Q 2, 4	2685.68 (6)	37223.5	37221.8	} 11.9	12.1
Q' 2, 4	2684.82 (3)	37235.4	—		
Q 1, 3	2683.11 (6)	37259.1	37257.7	} 11.8	11.6
Q' 1, 3	2682.26 (3)*	37270.9	—		
Q 0, 2	2681.10 (4)	37287.1	37286.8	} 11.2	11.3
Q' 0, 2	2680.29 (2)	37298.3	—		
Q 1, 2	2649.73 (4)	37728.5	37727.3	} 6.1	6.2
Q' 1, 2	2649.30 (2)	37734.6	—		
Q 0, 1	2647.50 (6)	37760.3	37760.3	} 5.9	5.6
Q' 0, 1	2647.08 (3)	37766.2	—		

\* There is a line or bandhead at  $\lambda 2682.19$  ( $\nu = 37272.1$ ) which is too strong to represent the isotope alone; the presence of this makes the observation of the true isotope difficult.

The differences between the calculated and observed positions of the bands may be attributed to imperfections of the formula as well as to errors of observation, and the agreements, both for positions and isotope displacements are too close to be accidental. It must be concluded that the bands degrading away from the red form part of the system which is mainly composed of bands degraded in the opposite direction.

It may be remarked that this feature is not quite identical with, for example, the occurrence in the CN spectrum of bands degraded to the red (the so-called "tail-bands" of CN) in the same parts of the spectrum as the chief bands which degrade away from the red. The possibility of this juxtaposition arises from the relatively small difference between the coefficients  $x'_0$ ,  $\omega'_0$ , and  $x''_0$ ,  $\omega''_0$ , so that high quantum numbers yield bands which appear in the same regions as those having smaller ones. In the present spectrum, this difference is also small, but the peculiar bands occur with *low* quantum numbers and have correspondingly great intensities.

As may be seen from Table I, the wave numbers along the sequences of the peculiar bands decrease in passing to the higher quantum numbers in a manner ordinarily regarded as appropriate for bands degraded towards the red, and thus provide an exception to the general rule.

#### *Separation of Q and R Heads.*

The spectrum of AlCl brings out very clearly some of the properties of band spectra, and for this reason it will be useful to collect the wave number differences between the Q and R heads in a form that will indicate the systematic character of the variations. This is done in Table IV.

Table IV.—Separations of R and Q Heads.

$v''$	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	61.8													
1	25.3	37.2												
2	12.6	17.3	25.8											
3		8.5	12.3	16.5	25.8									
4			7.5	9.2	11.9	16.2	23.1							
5				5.9	7.0	8.8	10.7	14.3	20.2					
6						5.6	6.6	7.5	9.3	12.0	15.2			
7								4.3	5.7	6.2	7.8	9.9	13.5	
8											5.0	5.0	5.5	
9												3.1	3.5	4.4

It will be observed that the RQ separations are comparatively large in the neighbourhood of the system origin and become small with the higher quantum numbers. The variations, however, are clearly systematic.

#### *The Isotope Bands.*

As already indicated, many bands due to the molecules  $\text{AlCl}^{37}$  are very clearly developed in association with the corresponding bands of the more

abundant molecules  $\text{AlCl}^{35}$ . Their wave numbers are included in the vibrational analysis, Table I, and in the general list of bands, Table VI.

It is of interest to compare the observed displacements of the bands of the heavier molecule with those calculated from theory. The isotopic displacement of a band  $v' v''$  is given by the relation\*

$$\begin{aligned} \nu_{v'v''}^t - \nu_{v'v''} = (\rho - 1) (\nu_{v'v''} - \nu_e) - (\rho - 1) [x'_e \omega'_e (v' + \tfrac{1}{2})^2 - x''_e \omega''_e (v'' + \tfrac{1}{2})^2] \\ - 2 (\rho - 1) [y'_e \omega'_e (v' + \tfrac{1}{2})^3 - y''_e \omega''_e (v'' + \tfrac{1}{2})^3], \end{aligned}$$

where  $\nu^t$  refers to the isotope and the various coefficients have the values already shown in the equation for the whole system of bands. The constant  $\rho$  is given by

$$\rho = \sqrt{\frac{\mu}{\mu^t}}$$

where  $\mu$  and  $\mu^t$  are the "reduced" or "effective" masses of the two molecules which are involved. If  $M_1$  and  $M_2$  be the masses of the two atoms in a molecule

$$\mu = \frac{M_1 M_2}{M_1 + M_2}.$$

For  $\text{AlCl}$  we have the masses:

$$\text{Cl}^{35} = 34.983; \text{Cl}^{37} = 36.980; \text{ and } \text{Al} = 26.97,$$

on the scale of  $\text{O} = 16$ , and with the result that  $\rho = 0.98818$ ; or  $(\rho - 1) = -0.01182$ .

With this value of  $(\rho - 1)$  the isotope displacements in  $\text{AlCl}$  are given by the equation

$$\begin{aligned} \nu_{v'v''}^t - \nu_{v'v''} = (\rho - 1) (\nu_{v'v''} - 38254.0) - (\rho - 1) [4.37 (v' + \tfrac{1}{2})^2 - 1.45 (v'' + \tfrac{1}{2})^2] \\ - 2 (\rho - 1) [0.216 (v' + \tfrac{1}{2})^3]. \end{aligned}$$

The displacement is towards higher or lower wave numbers according as  $(\nu^t - \nu)$  is positive or negative. The band due to the heavier molecule is always displaced from that of the lighter one in the direction of the system origin.

The observational data are collected in Table V, and the calculated values are enclosed in brackets for comparison.

The calculated values shown in Table V are vibrational displacements. Strictly, these should be corrected for rotational displacement for proper comparison with the observed values. The Q heads, being at or close to the band origins, are unaffected by the rotational effect, but the R' heads are

\* "Jevons' Report," p. 213.

Table V.—Isotope Displacements.

$\nu$	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	— (0.2)	5.9Q (5.8)	11.2Q (11.3)											
1	-5.4R (-4.9)		6.1Q (6.2)	11.8Q (11.6)										
2	-9.2 (-9.7)	-3.8R (-4.1)	1.3Q (1.4)	—	11.9Q (12.1)									
3		-8.8 (-8.6)	-3.1R (-3.1)	2.1 (2.3)	7.9 (7.6)									
4			-7.6 (-7.1)	-1.5R (-1.8)	3.8 (3.5)	8.6 (8.8)	13.8 (13.9)							
5				-5.8 (-5.6)		5.1 (5.1)	10.4 (10.2)	14.6 (15.2)	20.2 (20.2)	24.5Q (25.0)				
6							6.0 (6.9)	11.7R (11.9)	16.9 (16.8)	21.2 (21.8)	25.4 (26.6)	32.1R (31.2)		
7								10.0R (9.2)	13.9 (14.2)	19.3 (19.0)	23.8R (23.8)	28.6 (28.4)	32.2 (33.0)	
8											21.8R (21.6)	26.8R (26.2)	30.5R (30.9)	
9													28.7R (29.3)	31.6R (33.8)

displaced towards the origins of the respective bands by amounts nearly proportional to their distances from their origins. The amount of displacement is given approximately by  $2(\rho - 1) \times (R - Q)$ .<sup>\*</sup> For the AlCl bands, the displacement is approximately one-fortieth of the  $R - Q$  separations shown in Table IV. For bands on the less refrangible side of the system origin, the corrections are to be subtracted from the calculated vibrational displacements and added numerically for bands on the more refrangible side. As will be seen from Table IV, these corrections will not exceed 0.6 of a unit of wave number, except for the 0, 0 and 1, 1 bands in which the displacements are too small for measurement. This corresponds to about 0.04 Å. and is probably not appreciably greater than the errors of measurement of the distances between AlCl<sup>35</sup> and AlCl<sup>37</sup> heads. Thus the observed isotope separations for the R heads may be compared with the calculated vibrational displacements without serious error.

The observed separations entered in Table V are the means of measured intervals Q'Q, R'R when both have been observed; when only one has been measured the tabulated  $\Delta v$  is followed by the letter Q or R as the case may be.

It will be seen that the agreement between the calculated and observed displacements is very close for nearly all the bands. The difference exceeds one unit in only two of the bands, namely, R 6, 10 and R 9, 13 each of which is difficult to measure on account of low intensities. The mean difference between the observed and calculated values taken without regard to sign is 0.5 unit of wave number.

#### *Distribution of Intensities.*

The intensities of the bandheads have been estimated visually on a scale ranging from 6 for the strongest bands to 0 and 00 for the weakest. In fig. 1 these visual estimates are arranged in a  $v'v''$  scheme, with intensities substituted for wave numbers, and it will be seen that the locus of the stronger bands is in accordance with Condon's theoretical interpretation. As expected from the theory, since  $\omega'_0$  and  $\omega''_0$  are not very different, the parabola is a narrow one, as for the yellow bands of SrF and the green bands of BaF, the values of  $\omega'_0$  and  $\omega''_0$  being nearly the same for all the three spectra. The 0, 0 band is found to lie near the vertex of the curve and in following the sequence to higher values of  $v'$ ,  $v''$  the intensity is found to fall off in a regular manner.

<sup>\*</sup> See "Jevons' Report," p. 216, for exact formula.

Usually there are two intensity maxima separated by a minimum in accordance with the prediction of the theory that there are in general two preferred vibrational levels  $v''$  in the final electronic state for a given vibrational level  $v'$  in the initial electronic state.

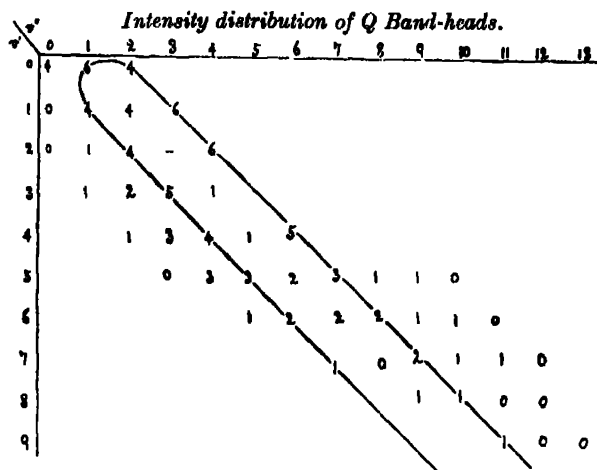


FIG. 1.

In general, the estimates of the relative intensities of the heads representing the more and less abundant molecules,  $\text{AlCl}^{35}$  and  $\text{AlCl}^{37}$ , are in accordance with the known ratio of the abundance of the chlorine isotopes (3 : 1). In a few bands the estimates were difficult on account of their superposition on the shading of other bands.

#### Classified Bands.

The band heads which have been classified are collected in order of wavelength in Table VI, and most of them will be readily identified in the photographs reproduced in Plate 8.

Table VI.—Classified Bands.

$\lambda\lambda$ (Intensity).	$\nu\nu$ .	Classification.	$\lambda\lambda$ (Intensity).	$\nu\nu$ .	Classification.
2809.57 (0)	35582.2	Q 7, 12	2782.57 (1)	35927.4	R 8, 12
08.50 (0)	595.7	R 7, 12	80.21 (0)	957.9	R 8, 12
07.08 (0)	613.8	Q 7, 12	75.64 (1)	36017.1	Q 7, 11
05.92 (00)	628.5	R 7, 12	74.88 (1)	027.0	R 7, 11
03.37 (0)	660.9	Q 6, 11	73.42 (0)	045.9	Q 7, 11
02.32 (0)	674.2	R 6, 11	72.69 (0)	055.4	R 7, 11
2799.80 (00)	706.3	R 6, 11	69.33 (1)	099.2	Q 6, 10
98.28 (0)	725.7	Q 5, 10	68.16 (0)	114.4	R 6, 10
91.94 (0)	806.8	Q 9, 13	67.43 (0)	124.0	Q 6, 10
91.60 (0)	811.2	R 9, 13	66.17 (00)	140.4	R 6, 10
89.14 (0)	842.8	R 9, 13	64.21 (1)	166.0	Q 5, 9
83.10 (0)	920.6	Q 8, 12	62.34 (0)	190.5	R, Q 5, 9

Table VI—(continued).

$\lambda\lambda$ (Intensity).	$\nu\nu$ .	Classification.	$\lambda\lambda$ (Intensity).	$\nu\nu$ .	Classification.
2758.76 (0)	36237.5	Q 9, 12	2653.68 (0)	37672.3	R <sup>+</sup> 3, 4
58.49 (1)	241.0	R 9, 12	49.73 (4)	728.5	Q 1, 2*
56.31 (0)	269.7	R 9, 12	49.30 (2)	734.6	Q <sup>+</sup> 1, 2*
49.83 (0)	355.2	Q 8, 11	47.50 (6)	760.3	Q 0, 1*
49.45 (1)	360.2	R 8, 11	47.08 (3)	766.2	Q 0, 1*
47.42 (0)	387.0	R <sup>+</sup> 8, 11	45.25 (1)	792.4	Q 7, 7
42.28 (1)	455.3	Q 7, 10	44.92 (2)	797.1	R 7, 7
41.69 (1)	463.1	R 7, 10	44.22 (1)	807.1	R <sup>+</sup> 7, 7
39.90 (0)	486.0	R <sup>+</sup> 7, 10	38.55 (2)	888.3	Q 6, 6
35.92 (1)	540.0	Q 6, 9	38.10 (3)	894.8	Q <sup>+</sup> R 6, 6
35.01 (1)	552.0	R 6, 9	37.59 (1)	902.1	R <sup>+</sup> 6, 6
34.35 (1)	561.0	Q <sup>+</sup> 6, 9	32.80 (3)	971.1	Q 5, 5
33.43 (0)	573.3	R <sup>+</sup> 6, 9	32.48 (0)	976.0	Q <sup>+</sup> 5, 5
30.58 (1)	611.4	Q 5, 8	32.19 (3)	979.9	R 5, 5
29.07 (1)	631.6	Q <sup>+</sup> R 5, 8	31.83 (1)	985.1	R <sup>+</sup> 5, 5
27.57 (0)	651.8	R <sup>+</sup> 5, 8	27.82 (4)	38043.0	Q 4, 4
26.06 (1)	672.1	Q 9, 11	27.55 (1)	046.9	Q 4, 4
25.83 (1)	675.2	R 9, 11	27.00 (3)	054.9	R 4, 4
17.10 (1)	703.1	Q 8, 10	26.74 (2)	058.7	R <sup>+</sup> 4, 4
16.73 (1)	798.1	R 8, 10	23.54 (5)	105.1	Q 3, 3
15.11 (0)	819.9	R <sup>+</sup> 8, 10	23.41 (2)	107.0	Q <sup>+</sup> 3, 3
09.41 (2)	897.5	Q 7, 9	22.40 (4)	121.6	R 3, 3
08.95 (2)	903.7	R 7, 9	22.24 (2)	124.0	R <sup>+</sup> 3, 3
07.97 (0)	917.1	Q <sup>+</sup> 7, 9	19.98 (4)	156.9	Q 2, 2
07.56 (1)	922.7	R <sup>+</sup> 7, 9	19.89 (2)	158.2	Q <sup>+</sup> 2, 2
02.94 (2)	985.8	Q 6, 8	18.21 (3)	182.7	R 2, 2
02.26 (3)	995.1	R 6, 8	16.97 (4)	200.5	Q 1, 1
01.70 (1)	37002.8	Q <sup>+</sup> 6, 8	14.44 (8)	237.7	Q 0, 0, and R 1, 1
01.04 (1)	011.8	R <sup>+</sup> 6, 8			
2697.43 (3)	061.3	Q 5, 7	10 22 (5)	299.5	R 0, 0
96.39 (3)	075.6	Q <sup>+</sup> R 5, 7	07.06 (1)	345.9	Q 0, 5
95.31 (1)	090.5	R <sup>+</sup> 5, 7	06.68 (2)	351.5	R 6, 5
92.80 (5)	125.1	Q 4, 6	01.14 (3)	433.2	Q 5, 4
91.80 (2)	138.9	Q <sup>+</sup> 4, 6	00.67 (3)	440.2	R 5, 4
91.12 (2)	148.2	R 4, 6	2596.04 (3)	508.7	Q 4, 3
90.13 (1)	161.9	R <sup>+</sup> 4, 6	95.52 (1)	516.4	R <sup>+</sup> 4, 3
85.68 (6)	223.5	Q 2, 4*	95.42 (2)	517.9	R 4, 3
85.05 (1)	232.2	Q 8, 9	91.64 (2)	574.1	Q 3, 2
84.82 (3)	235.4	Q <sup>+</sup> 2, 4*	91.02 (1)	583.3	R <sup>+</sup> 3, 2
84.56 (1)	239.0	R 8, 9	90.81 (2)	586.4	R 3, 2
83.11 (6)	259.1	Q 1, 3*	87.88 (1)	630.1	Q 2, 1
82.26 (3)	270.9	Q <sup>+</sup> 1, 3*	86.98 (1)	643.6	R <sup>+</sup> 2, 1
81.10 (4)	287.1	Q 0, 2*	86.72 (2)	647.4	R 2, 1
80.29 (2)	298.3	Q <sup>+</sup> 0, 2*	84.70 (0)	677.6	Q <sup>+</sup> 1, 0
77.12 (0)	342.5	Q 7, 8	83.37 (0)	697.5	R <sup>+</sup> 1, 0
76.71 (2)	348.2	R 7, 8	83.01 (0)	702.9	R 1, 0
76.13 (0)	356.3	Q <sup>+</sup> 7, 8	75.71 (0)	812.2	R 6, 4
75.70 (1)	362.3	R <sup>+</sup> 7, 8	70.40 (0)	892.8	Q <sup>+</sup> 5, 3
70.47 (2)	435.5	Q 6, 7	70.02 (0)	898.6	Q, R <sup>+</sup> 5, 3
69.93 (2)	443.0	R 6, 7	69.63 (1)	904.5	R 5, 3
69.10 (1)	454.7	R <sup>+</sup> 6, 7	65.27 (0)	970.6	Q <sup>+</sup> 4, 2
64.84 (2)	514.6	Q 5, 6	64.76 (1)	978.3	Q, R <sup>+</sup> 4, 2
64.08 (2)	525.3	Q <sup>+</sup> R 5, 6	64.27 (1)	985.8	R 4, 2
63.35 (0)	535.5	R <sup>+</sup> 5, 6	60.75 (0)	39039.4	Q <sup>+</sup> 3, 1
60.06 (1)	581.9	Q 4, 5	60.16 (1)	048.4	Q, R <sup>+</sup> 3, 1
59.46 (0)	590.4	Q <sup>+</sup> 4, 5	59.60 (1)	056.9	R 3, 1
58.92 (2)	598.1	R 4, 5	56.87 (0)	098.6	Q <sup>+</sup> 2, 0
58.31 (1)	606.7	R <sup>+</sup> 4, 5	56.28 (0)	107.6	Q 2, 0
56.03 (1)	639.0	Q 3, 4	56.07 (0)	110.8	R <sup>+</sup> 2, 0
55.45 (0)	647.2	Q <sup>+</sup> 3, 4	55.46 (1)	120.2	R 2, 0
54.21 (1)	664.8	R 3, 4			

\* Degraded away from red.

For completeness, particulars of a few faint bands which do not appear in the list of classified bands are given in Table VII. Some of these may possibly represent lines, as it was very difficult to distinguish faint band heads from faint lines unless places could be found for them in the vibrational analysis.

In addition to the bands, a few distinct lines also occur in the photographs of the aluminium chloride spectrum. The most prominent are due to aluminium (Al II) including  $\lambda\lambda$  2816.2 (4) 2660.39 (2), 2652.48 (5), 2631.60 (1), 2575.40 (2),

Table VII.—Faint Unclassified Bands.

$\lambda\lambda$ Int.	$\nu\nu$ .	$\lambda\lambda$ Int.	$\nu\nu$ .
2810.12 (0)	35875.2	2759.47 (0)	36228.1
2788.19 (0)	855.0	2721.82 (1)	729.3
2770.88 (0)	962.2	2719.95 (1)	754.6
2777.90 (0)	987.8	2706.53 (0)	936.7
2777.48 (0)	993.3	2689.72 (0)	37167.6
2766.78 (0)	36132.5	2686.13 (1)	217.3

2575.12 (3) and 2568.00 (3). Other lines which appear with varying intensities, but are mostly faint, are due to silicon, from disintegration of the discharge tube, and to iron from a slight impurity in the aluminium electrodes. There are still other lines which are probably due to Cl II, namely,  $\lambda\lambda$  2687.95 (3),\* 2620.19 (1), 2617.02 (4) and 2614.32 (3). Unidentified lines include strong lines at  $\lambda\lambda$  2688.90 (5),\* 2614.19 (5) and fainter ones at  $\lambda\lambda$  2670.58 (1), 2636.44 (2).

\* [Added in proof May 1, 1934—Professor W. E. Curtis has pointed out to the authors that the two "lines" recorded at  $\lambda\lambda$  2688.90(5), 2687.95(3) (see Plate 8, b) closely agree in positions with the Q and Q' heads which might be expected at  $\nu'\nu''$  — (3, 5). The wave numbers are respectively 37178.9 and 37192.0, while the position of the Q head calculated from the general formula is 37178.8 and the difference  $\Delta\nu = 13.1$  compares favourably, with the calculated isotope separation 12.4; the relative intensities are also appropriate to Q and Q'. It is therefore very probable that Q (3, 5) marks a point at which the B values\* for the upper (B') and lower (B'') states are so nearly equal that the band lines of the Q and Q' branches are superposed and thus present the appearance of simple lines or undegraded bands.

With this addition the  $\nu'' - \nu' = 2$  sequence is complete from (0, 2) to (9, 11). From (4, 6) to (9, 11) the degradation is towards the red, and from (0, 2) to (2, 4) away from the red, (3, 5) being a transition stage. There must accordingly be a "crossing over" of the B values in this sequence near the (3, 5) bands; that is, for pairs of values of  $\nu'\nu''$  greater

\* In the rotational term,  $B = h/8\pi^2cI$ , where I is the moment of inertia of the rotating molecule, and the other symbols have their usual meanings.



*Summary.*

A band system having its greatest intensity at  $\lambda$  2610 has been photographed with high dispersion in vacuum tubes containing the vapour of aluminium chloride. The system extends from  $\lambda$  2555 to about  $\lambda$  2810, and with a few important exceptions, the bands are degraded to the red. The heads of the bands occur in pairs which have been interpreted as representing R and Q branches. Along with the main bands due to the lighter molecules  $\text{AlCl}^{35}$  there are associated bands which originate in the heavier molecules  $\text{AlCl}^{37}$  and are displaced in accordance with theoretical expectation.

An unusual feature of the band system is the occurrence of two short sequences of bands which are degraded away from the red, in opposition to the remaining bands. These have low vibrational quantum numbers and are

than (3.5), ( $B' - B''$ ) must be negative, while for smaller values it must be positive. The observations also show that there is a similar crossing over in the  $v'' - v' = 1$  sequence.

In the absence of rotational analyses, the B values are not known, but Professor Curtis has suggested a procedure by which relative values of ( $B' - B''$ ) derived from separations of the R and Q heads in the "normal" bands may be extrapolated to give approximate values for the bands which are degraded in the opposite direction. In this way it can be shown that a change of sign of ( $B' - B''$ ) is to be expected in the positions where the bands degraded away from the red are observed.

Theory (Jevons' Report, p. 44) gives

$$v_R - v_Q = - \frac{(B' + B'')^2}{4(B' - B'')}$$

and since ( $B' + B''$ ) will vary relatively slowly with  $v$ , the ( $B' - B''$ ) values will be approximately proportional to the reciprocals of  $v_R - v_Q$ . Reciprocals of the separations of the normal R and Q heads given in Table IV were accordingly arranged in a  $v'v''$  table, and were found to show a definitely linear decrease of the negative values of ( $B' - B''$ ) with increasing  $v''$ , while they increased with  $v'$  more rapidly than a linear law would indicate. Extrapolation then gave positive values for (0.1), (0.2), (1.3), (2.4) and sensibly zero for (1.2). Small negative values were found for (2.3) and (3.5). The band (2, 3) is not recorded in the observations, but interpolation of its position suggests that the Q head is probably obscured by the strong aluminium line  $\lambda$  2652.48 ( $\nu$  37689.4), and re-examination of the photographs has shown that there is a faint line or band which may represent the Q head; comparison with other members of this sequence suggests that (2, 3) would not be a strong band.

Notwithstanding the low degree of accuracy of the ( $B' - B''$ ) values obtained in this way, these results provide a satisfactory interpretation of the reversed direction of degradation of some of the bands which appear in  $\text{AlCl}$ , and of the occurrence of undegraded bands. In the bands degraded away from the red, P heads should occur instead of R, but these have not been noted. They would be expected to appear at relatively large distances from the Q heads, and they may thus be observed by adjacent bands, or be too feeble for observation.]

of correspondingly high intensity. The close agreement with the positions calculated for them, together with the displacements of the accompanying isotope heads, shows that these bands are part of the same system as the bands which are degraded to the red.

The vibrational analysis and the discussion of the isotope bands combine to prove that the band system under investigation originates in the diatomic molecules  $\text{AlCl}^{36}$  and  $\text{AlCl}^{37}$ .

---

*Investigations in the Infra-red Region of the Spectrum. Part X.—  
The Asymmetrical Molecule Nitrosyl Chloride, NOCl.*

By C. R. BAILEY and A. B. D. CASSIE, the Sir William Ramsay Laboratories of Inorganic and Physical Chemistry, University College, London.

(Communicated by F. G. Donnan, F.R.S.—Received January 20, 1934.)

The infra-red absorption spectrum of nitrosyl chloride was investigated with two objects in view. Firstly, if the substance is a covalent compound with normal valency linkings, the molecule will probably be asymmetrical and triangular. If this is so, the absorption bands will show no pronounced branch maxima such as are predicted and observed for symmetrical molecules, and the presence or absence of these maxima in the same conditions of resolution and concentration provides a test of the theory according to which the spectra of such gases as  $\text{SO}_2$ ,  $\text{ClO}_2$ , and  $\text{Cl}_2\text{O}$  have been interpreted. The two former molecules are required by the theory to be symmetrical, although such symmetry is not to be expected from the ordinary rules of valence. The experimental fact that the characteristic P, Q, and R branch maxima are absent in NOCl thus provides a verification of the theory.

Secondly, six electrons are effectively available for binding in the molecule, and accordingly\* the structure should be triangular with a vertical angle ONCl approximately  $120^\circ$ , and governed by a valence force system. The molecule should, therefore, show whether or not the structure of triatomics is determined primarily by the number of available binding electrons.

*Experimental.*

The absorption spectrum of NOCl prepared by heating dry sodium chloride with nitrosylsulphonic acid shows that such specimens are contaminated with

\* Cassie, 'Nature,' vol. 131, p. 438 (1933).

$\text{N}_2\text{O}_4$ . The older method of direct union of nitric oxide and chlorine due to Gay Lussac as modified by Taylor and Denslow\* was adopted. Chlorine was obtained from a commercial cylinder and carefully dried, while nitric oxide was produced by mixing and warming saturated solutions of sodium nitrite and ferrous sulphate acidified with sulphuric acid. Snow, Rawlins, and Rideal† have shown that the infra-red spectrum of the gas prepared in this way is free from bands due to  $\text{NO}_2$ . It was passed successively through water, sodium hydroxide, concentrated sulphuric acid, and finally dried by phosphoric oxide.

The complete system from the NO generator to the final NOCl receiver was of glass. Oxygen was removed by passing nitrogen through the apparatus for some time before the sodium nitrite solution was dropped into the acidified ferrous sulphate. Chlorine was liquefied in a trap, and the NOCl formed there distilled across  $\text{P}_2\text{O}_5$  to a second trap where it was frozen, and any excess NO pumped off. The possible impurities are now chlorine and  $\text{NOCl}_2$ . The former has no infra-red spectrum, while a sample of NOCl on passing through a glass spiral heated to  $150^\circ$  gave all the bands previously obtained with unaltered relative intensities.

### Observations.

The absorption spectrum is given in Table I. No distinct maxima were found in any of the bands, and the usual P — R branch separation of the symmetrical molecules is replaced by an estimate of the band width.

Table I.

Band.	Band centre.		Band width. ( $\text{cm.}^{-1}$ ).	Intensity.	Slit width ( $\text{cm.}^{-1}$ ).
	$\lambda$ ( $\mu$ ).	$\nu$ ( $\text{cm.}^{-1}$ ).			
A	15.8	633	25	6	5
B	10.83	923	15	10	5
C	8.33	1200	20	3	12
D	5.46	1832	20	7	8
E	4.64	2155	25	4	10

*Band A at 15.8  $\mu$ .*—This region was particularly carefully investigated for P, Q, and R maxima as the longest wave-length band is usually most readily

\* 'J. Phys. Chem.,' vol. 31, p. 374 (1927).

† 'Proc. Roy. Soc.,' A, vol. 124, p. 453 (1929).

resolved. A peculiar contour was obtained with a very broad maximum near  $655\text{ cm.}^{-1}$ , and a smaller indistinct rise at  $627\text{ cm.}^{-1}$ . None of the bands showed any of the characteristics of normal resolution such as was obtained in COS, a molecule which has only a slightly smaller molecular weight, or in the even heavier molecule,  $\text{Cl}_2\text{O}$ .

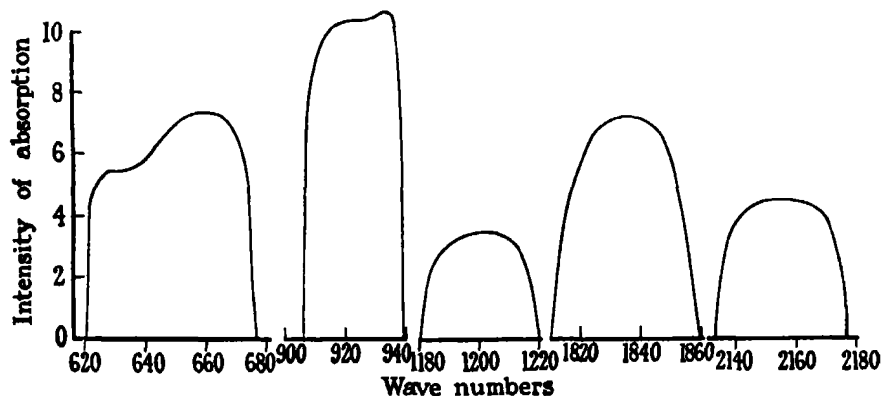


FIG. 1.—Infra-red absorption spectrum of NOCl.

*Band B at  $10.83\text{ }\mu$ .*—Two indistinct maxima were found some  $10\text{ cm.}^{-1}$  apart. The band is somewhat narrower than the others.

*Bands C, D, E, at  $8.33$ ,  $5.46$ , and  $4.64\text{ }\mu$ .*—Broad maxima with undulating contours were obtained in each case.

#### Discussion.

Lack of pronounced branch maxima is a great handicap in fixing the structure and force system of the molecule, since these features have previously been used in deciding between the acute and obtuse angled structures, and in differentiating between the two possible force systems. Furthermore, the determination of a third force constant reduces the possibility of giving a precise interpretation of the spectrum in terms of the molecular parameters. However, by the use of reasonable assumptions and slight approximations we have reached what seem to us satisfactory conclusions.

#### The Central Force System.

The normal frequencies for the asymmetric molecule in this system have been evaluated by Radakovic,\* and for the present purpose the relations of

\* 'Mhft. Chem.,' vol. 36, p. 555 (1930).

the roots to the coefficients of the frequency equations prove most useful. We have

$$p_1^2 + p_2^2 + p_3^2 = K_1/\mu_1 + K_2/\mu_2 + K_3/\mu_3; \quad (1)$$

$$\Sigma p_1^2 p_2^2 = \Sigma K_1 K_2 / \mu_1 \mu_2 - \Sigma K_1 K_3 (\cos^2 \alpha_1) / m_1^2; \quad (2)$$

$$p_1^2 p_2^2 p_3^2 = K_1 K_2 K_3 \left\{ 1 / \mu_1 \mu_2 \mu_3 - \Sigma (\cos^2 \alpha_3) / m_3^2 \mu_1 + \frac{2 (\cos \alpha_1 \cos \alpha_2 \cos \alpha_3)}{m_1 m_2 m_3} \right\}. \quad (3)$$

where  $p = 2\pi c\nu$ , and  $\nu$  is a fundamental frequency in wave numbers, while  $1/\mu_1 = 1/m_1 + 1/m_2$ ,  $1/\mu_2 = 1/m_2 + 1/m_3$ , and  $1/\mu_3 = 1/m_3 + 1/m_1$  (equation (1)).

The first equation shows that the sum of the fundamental frequencies is independent of the angles of the triangle, and that this sum is a measure of the force constants of the molecule. The first point to note in applying formula (1) is that if the unit mass be taken as 10 times that of the H atom, that is  $1.65 \times 10^{-23}$  gm.,  $1/\mu$  is of the order of unity. Hence if the left-hand side of (1) be multiplied by this factor, the equation immediately indicates possible values for the force constants in dynes/cm. On assuming the most intense bands A, B, and D of Table I to give the fundamental frequencies, the left-hand side of (1) becomes  $27.2 \times 10^5$  dynes/cm., and possible values of the force constants might be  $K_1 = 9.8$ ,  $K_2 = 9.8$ , and  $K_3 = 4 \times 10^5$  dynes/cm., or  $K_1 = 13$ ,  $K_2 = 6.5$ , and  $K_3 = 5 \times 10^5$  dynes/cm. The first solution would indicate a molecule like  $\text{SO}_2$  with symmetrical bonds, and the second a double bond between N and O, and a single bond between N and Cl. Both structures may be considered in dealing with equations (2) and (3), and in taking  $K_1 = K_2$ , or  $K_1 = 2K_2$  we effectively reduce the unknown force constants from three to two. Insertion of possible arithmetical values in (2) indicates that the terms involving the angles have but approximately one-tenth of the value of the other terms and as a first approximation may be neglected. From (1) and (2), if  $K_1 = \gamma K_2$ , then

$$(\gamma^2/\mu_2^2 + \gamma/\mu_1\mu_2 + 1/\mu_1^2) K_1^2 - \Sigma p_1^2 (1/\mu_1 + \gamma/\mu_2) K_1 + \Sigma p_1^2 p_2^2 = 0. \quad (4)$$

I (a).  $\gamma = 1$  gives  $K_1 = 2.8 \times 10^5$  or  $12.3 \times 10^5$  dynes/cm. The first solution may obviously be discarded, while the second gives a negative value for  $K_3$ .

I (b).  $\gamma = \frac{1}{2}$  gives  $K_1 = 4.1$  or  $14 \times 10^5$  dynes/cm. The first solution is impossible, while the second gives  $K_3$  as  $1.3 \times 10^5$  dynes/cm., a not unlikely value. Insertion of the corresponding values in equation (3), even with the

most favourable values of the angles, shows a discrepancy of 30% between the left- and right-hand sides of the equation. This fact in itself does not absolutely discredit the central force system, for the actual force constants may lie between the limits  $\gamma = 1$  and  $\gamma = \frac{1}{2}$ , but we shall find that consistent agreement between all data is provided by the valence force system, and the latter is consequently to be preferred.

*Valence Force System.*

The normal frequencies for this system in the same conditions have been determined by Lechner,\* and again we shall use the relations of the roots to the coefficients of the frequency equation. They are :

$$\Sigma p_1^2 = K_1/\mu_1 + K_2/\mu_2 + d/\mu_4; \quad (5)$$

$$\begin{aligned} \Sigma p_1^2 p_2^2 = & (K_1 K_2 / \mu_1 \mu_2) [1 - (\mu_1 \mu_2 / m_2) \cos^2 \theta] \\ & + (K_1 d / \mu_1 \mu_4) [1 - (\mu_1 \mu_4 / m_2^2) (s/s_{23})^2 \sin^2 \theta] \\ & + (K_2 d / \mu_2 \mu_4) [1 - (\mu_1 \mu_4 / m_2^2) (s/s_{12})^2 \sin^2 \theta]; \end{aligned} \quad (6)$$

and

$$p_1^2 p_2^2 p_3^2 = (K_1 K_2 d / \mu_1 \mu_2 \mu_4) [1 - \mu_1 \mu_2 / m_2^2]. \quad (7)$$

$\theta$  is the complement of  $\alpha$  (equation (2));  $s$ ,  $s_{12}$ , and  $s_{23}$  are such that the potential energy of a displacement  $\Delta\theta$  is given by

$$V = \frac{1}{2} d_{12} (s_{12} \Delta\theta)^2 = \frac{1}{2} d (s \Delta\theta)^2 = \frac{1}{2} d_{23} (s_{23} \Delta\theta)^2, \quad (8)$$

where  $d_{12}$ ,  $d$ , and  $d_{23}$  are tangential restoring force constants, and together with  $s_{12}$ ,  $s$ , and  $s_{23}$  take into account the unequal sides of the triangle.  $d$  may be taken as equal to  $K/4$ , where  $K$  is the tangential restoring force for the symmetrical triangle. Finally

$$1/\mu_4 = (1/\mu_1) (s/s_{12})^2 + (2s^2/m_2 s_{12} s_{23}) \cos \theta + (1/\mu_2) (s/s_{23})^2.$$

Here again we investigate the limiting cases  $K_1 = K_2$ , and  $K_1 = 2K_2$ . Equations (5) and (7) are clearly the equations to take in solving for the unknown force constants, since  $\theta$  only appears in these through  $\mu_4$ . The value of  $1/\mu_4$  is  $3.8/10m_H$  when  $\theta = 0$ ,  $3.1/10m_H$  when  $\theta = 60^\circ$ , and  $2.4/10m_H$  when  $\theta = 90^\circ$ , and so does not vary greatly with  $\theta$ . Lack of branch maxima shows that the molecule is not linear like COS, and the most probable configuration has  $\theta = 60^\circ$ , and the corresponding value of  $1/\mu_4$  will be used. The effective radii of the neutral atoms are taken as N— = 0.61, O— = 0.60, N— = 0.71, and Cl— =  $0.97 \times 10^{-8}$  cm.  $s_{12}$ ,  $s_{23}$ , and  $s$  will then be taken as 1.2, 1.7, and 1.4 Å. Equation (5) now gives a measure of the force con-

\* 'SitzBer. Akad. Wiss., Wien,' vol. 141, p. 291 (1932).

stants equivalent to that given by (3) for the central force system. Eliminating  $K_2$  and  $d/\mu_4$  from (5), (7), and the relation  $K_1 = \gamma K_2$ , we have

$$\gamma (1/\mu_1 + \gamma/\mu_2) K_1^3 - (\gamma \Sigma p_1^2) K_1^2 + (p_1^2 p_2^2 p_3^2)/(1/\mu_1 \mu_2 - 1/m_2^2) = 0, \quad (9)$$

from which  $K_1$  may be determined. Equation (9) does not contain the uncertain quantity  $1/\mu_4$ .

II (a).  $\gamma = 1$ . The possible solution is  $K_1 = K_2 = 10.2$ ,  $K_3 = 4 \times 10^5$  dynes per centimetre.

II (b).  $\gamma = \frac{1}{2}$ .  $K_1 = 13$ ,  $K_2 = 6.5$ , and  $K_3 = 4 \times 10^5$  dynes/cm. A choice of either of these two possible solutions may be made on investigation of equation (6). This involves only one quantity,  $\theta$ , which is now unknown, and the solutions (a) and (b) are tested by solving (6) for  $\theta$ . In this respect we may mention that we have tested other relations of  $K_1$  to  $K_2$ , such as  $\gamma = 1/3$ , but the resulting agreement is poor. A negative value for  $\sin^2 \theta$  is obtained from case II (a) while case II (b) gives  $\theta$  as approximately  $35^\circ$ . Hence the valence force system with a vertical angle of some  $140^\circ$  most satisfactorily represents the characteristics of the molecule.

### Band Assignments.

An unambiguous allocation seems impossible for the asymmetric molecule, even if the Raman spectrum were known. We have consequently assigned the vibration  $\nu_3$ , which in the symmetrical triangle corresponds to the asymmetric mode and is generally the most intense in the infra-red, to band B at  $923 \text{ cm.}^{-1}$ , since this has the greatest intensity and least width. On this assumption the remainder of the bands are allotted as in Table II.

Table II.—Allocation of bands observed in the spectrum of NOCl.

Band.	Centre ( $\text{cm.}^{-1}$ ).	Mode.	Calculated for ( $\text{cm.}^{-1}$ ).	Difference.
A	633	$\nu_1$	—	—
B	923	$\nu_3$	—	—
C	1200	$2\nu_2$	1266	-66
D	1832	$\nu_1$	—	—
E	2155	$2\nu_2 + \nu_3$	2189	-34

Alternative possibilities are: in the case of band C,  $\nu_1 - \nu_2$  is calculated as  $1199 \text{ cm.}^{-1}$ , but in spite of the nearness of this value to that observed it

seems preferable to adopt the interpretation given in Table II, since difference tones are likely to be weak; band E is capable of at least two other interpretations, and indeed may be due to a superposition of the three modes. The latter band has a comparatively high intensity, which might be further associated with the energy of activation of the NOCl molecule, which Taylor and Denslow (*loc. cit.*) have shown to correspond to  $4.77 \mu$ , and some form of resonance may anomalously increase the intensity of this combination tone.

### *The Molecular Structure.*

The structure and force constants obtained are of considerable interest. In spite of the asymmetry of the molecule, the six valence electrons available for binding place the molecule in the  $\text{SO}_2$  class\* having a triangular structure with a vertical angle of some  $120^\circ$ , satisfying the valence force system, and with a tangential restoring force constant of approximately  $3.5 \times 10^5$  dynes/cm. Furthermore, the force constant total is the same for both  $\text{SO}_2$  and NOCl, being  $19.5 \times 10^5$  dynes/cm. and the six binding electrons appear to make the same contribution in both.

The heat of the reaction  $2(\text{NO}) + (\text{Cl}_2) = 2\text{NOCl}$  has been found by Dixon† to be 18.6 k. cal. Accordingly  $(\text{N}) + (\text{O}) + (\text{Cl}) = 188$  k. cal., and allotting this value to the links in proportion to the force constants, we obtain 125 k. cal. for the  $\text{N} = \text{O}$  bond, and 63 k. cal. for  $\text{N} - \text{Cl}$ . This is in agreement with the structure of the nitrous oxide molecule which we have shown to be  $\text{N} = \text{N} = \text{O}$ , with force constants  $K_1 = K_2 = 14 \times 10^5$  dynes per centimetre, and a heat of formation of 256 k. cal. Again the force constant‡ for nitric oxide, NO, is  $15.7 \times 10^5$  dynes/cm., the heat of formation is 150 k. cal., and we have for the  $\text{N} = \text{O}$  bond in NOCl,  $150 \times 13/15.7 = 124$  k. cal. in excellent agreement. The value of the  $\text{N} - \text{Cl}$  link in nitrogen trichloride seems to be some 45 k. cal., the smaller heat being presumably associated with the instability of this substance.

### *Raman and other Spectra.*

Dadiou and Kohlrausch§ were unable to obtain a Raman spectrum from liquid NOCl, and assumed that the liquid was too dark. It is probable that

\* Cassie, *loc. cit.*

† 'Z. phys. Chem.,' Bodenstein Festband, p. 679 (1931).

‡ Bailey and Cassie, 'Phys. Rev.,' vol. 39, p. 534 (1932).

§ 'Phys. Z.,' vol. 33, p. 165 (1932).



the asymmetry of the molecule results in comparatively weak displacements. The absorption spectrum in the visible and ultra-violet has been examined in part by Magnanini,\* Kistiakowsky† and Leermakers and Ramsperger,‡ but apart from the recognition of a banded structure with partial resolution, no quantitative results are available. The photochemical reaction was interpreted by Kistiakowsky as involving  $\text{NOCl}' + \text{NOCl}$  in spite of the fact that a large reduction in pressure, and the seven-fold addition of nitrogen had no effect on the quantum efficiency,  $\gamma$  retaining the value of 2 independent of  $\lambda$ . The reason given for the assumption was that the heat of dissociation of  $\text{NOCl}$  into  $\text{NO} + \text{Cl}_2$  corresponded to  $613 \mu\mu$ , whereas light of longer wave-length was effective; the calculated wave-length is incorrect as a consequence of the belief that the heat of formation of one mol. of  $\text{NOCl}$  from  $\text{NO} + \text{Cl}_2$  was 18 k. cals., the latter value, however, corresponding to two mols. The complete quantum yield is thus more likely to be explained by primary dissociation, and the suggestions made by Allmand§ are probably correct.

The most reasonable values for the molecular constants of nitrosyl chloride may be summarized as in Table III.

Table III.—The Molecular Constants of  $\text{NOCl}$ .

Frequencies ( $\text{cm}^{-1}$ ).	Force constants (dynes/cm. $\times 10^{-8}$ ).	Dimensions.
$\nu_1 = 1832$ $\nu_2 = 633$ $\nu_3 = 923$	$K_1 = 13 \text{ (N = O)}$ $K_2 = 6.5 \text{ (N - Cl)}$ $K_\theta = 4$	$\widehat{\text{ONCl}} = \text{ca. } 140^\circ$ $\text{N - Cl} = 1.7 \text{ \AA.}$ $\text{N = O} = 1.2 \text{ \AA.}$

Taking  $\nu_2$  as  $a_2 + b_2$ , where  $b_2$  from band C ( $2a_2 + 4b_2$ ) is  $-33 \text{ cm}^{-1}$ , we can calculate approximately the heat of dissociation corresponding to this vibration from the formula  $D = a^2/4b$ . The value is some 33 k. cals. which may reasonably correspond to the reaction  $\text{NOCl} = \text{NO} + \text{Cl} - 38 \text{ k. cal.}$

The authors once again thank Professor F. G. Donnan, C.B.E., F.R.S.; and the Department of Scientific and Industrial Research for a Senior Award to A. B. D. C.

\* 'Z. phys. Chem.,' vol. 4, p. 427 (1889).

† 'J. Amer. Chem. Soc.,' vol. 52, p. 102 (1930).

‡ *Ibid.*, vol. 54, p. 1837 (1932).

§ 'Chem. Soc. Discussion,' p. 34 (1931).

The infra-red absorption spectrum of nitrosyl chloride, NOCl, has been examined from 1 to 18  $\mu$ . Five bands were isolated between 4 and 16  $\mu$ , but as a consequence of the expected asymmetry of the molecule, no resolution was obtained. The molecule falls into the sulphur dioxide class, having a vertical angle of some  $140^\circ$ , and obeying a valence force system. The force constants have been determined and associated with the heats of linking.

### *The Quantum Theory of the Neutron.*

By G. TEMPLE, Ph.D., D.Sc., University of London, King's College.

(Communicated by Sir Arthur Eddington, F.R.S.—Received January 29, 1934.)

#### 1. *Introduction.*

The object of this paper is to show that a plausible theory of the neutron can be developed from Dirac's wave equation without the use of any *ad hoc* assumptions. It is shown that the second order wave equation of the hydrogen atom, which exhibits the relativistic and spin corrections, possesses two sets of solutions "H" and "N" distinguished by their behaviour as  $r \rightarrow 0$  ( $r$  being the distance of the electron from the proton).

The H-solutions are the accepted wave functions of the hydrogen atom. As  $r \rightarrow 0$  these solutions tend to zero if the serial quantum number  $l$  differs from zero, and they become infinite of order  $r^{(1-\alpha)^{1/2}-1}$  if  $l = 0$  ( $\alpha$  is the fine structure constant).

The energy levels in the discrete spectrum are given approximately by the formula

$$E = m_0 c^2 (1 - \alpha^2/2n^2), \quad (n = 1, 2, 3, \dots),$$

where  $m_0 c^2$  is the proper energy of the electron. The energy of the ground state is

$$E_H = m_0 c^2 (1 - \frac{1}{2}\alpha^2).$$

The N-solutions exist only if  $l = 0$ . As  $r \rightarrow 0$  they become infinite of order  $r^{-(1-\alpha)^{1/2}}$ , but, like the H-solutions, they are quadratically integrable over the range,  $0 \leq r \rightarrow \infty$ . The energy levels of the N-solutions form a continuous spectrum, the values of the energy  $E$  covering the half open interval,

$$-m_0 c^2 \leq E < +m_0 c^2,$$

and the lowest value of the energy being

$$E_N = -m_0c^2.$$

It is suggested here that the N-solutions are the wave functions of the neutron. On making this identification the theory attributes the following properties to the neutron :—

(a) The binding energy of the neutron in its normal state is

$$E_H - E_N = m_0c^2 (2 - \frac{1}{2}\alpha^2),$$

i.e., approximately  $2m_0c^2$  or  $10^6$  electron volts. Thus the mass of the neutron is less than the mass of the proton by the amount of twice the mass of an electron. This is in agreement with the estimate made by Chadwick in his Bakerian lecture.\*

(b) The mean value of  $r^{-1}$  for the neutron in its normal state is

$$\overline{1/r} = 6/\alpha_0^2 a_0,$$

where  $a_0$  is the radius of the first Bohr orbit of hydrogen, i.e., in the usual notation,  $a_0 = \hbar^2/(m_0c^2), \dagger$

The mean value of  $r$  is  $\bar{r} = 3a_0/10$ .

(c) At a large distance  $r$  from a neutron in its normal state the electrostatic potential of the field is given approximately by the equation,

$$V(r) = \frac{3\sqrt{2}}{16} \cdot \frac{e}{(a_0r)^3} \exp \{-2(8r/a_0)^{\frac{1}{2}}\}.$$

(d) The serial quantum number of the neutron is  $l = 0$ . Hence the charge of the neutron is spherically symmetrical and the electron in it has a spin of  $\frac{1}{2}\hbar$ . Hence the total spin of the neutron must be  $\hbar$  or zero. At present it is premature to say whether this deduction is or is not in agreement with experiment.

(e) It remains to consider the probabilities of transitions from a hydrogen atom to a neutron. Since  $l = 0$  for a neutron, such transitions can occur only when  $l = 1$  for the hydrogen atom. Hence the normal state of the hydrogen atom ( $n = 1, l = 0$ ) is a meta-stable state, which cannot collapse into a neutron. This result is satisfactory as far as it goes, but it is clearly far from a complete explanation of the rarity of the transitions from hydrogen atoms to neutrons.

\* 'Proc. Roy. Soc.,' A, vol. 142, p. 1 (1933).

†  $\hbar$  is Planck's constant divided by  $2\pi$ .

The preceding deductions from the theory of this paper indicate that although it requires some modification and extension, it does succeed in reproducing the main properties attributed to neutrons. The strength of the theory lies in the fact that it is free from any special assumptions, all the results being deduced from the second order wave equation for the hydrogen atom.

## 2. The First Order Wave Equation.

The first order wave equation for the hydrogen atom may be written as

$$F\psi \equiv \{(p_1\sigma_1 + p_2\sigma_2 + p_3\sigma_3) \rho_1 + ic^{-1}(E + e^2/r) \rho_3\} \psi = \pm im_0c\psi, \quad (2.1)$$

where  $\rho_1, \rho_2, \rho_3$  and  $\sigma_1, \sigma_2, \sigma_3$  are two commuting sets of quaternion operators. It is convenient to abbreviate such a quasi-scalar product as

$$(p_1\sigma_1 + p_2\sigma_2 + p_3\sigma_3)$$

to  $(p, \sigma)$ .

The solution of this wave equation depends upon the identity,

$$(x, \sigma)(p, \sigma) \equiv rp_r + i\rho_3\kappa,$$

where  $p_r$  is the radial momentum operator,

$$p_r = \frac{1}{2} \{(x_1/r) p_1 + p_1 (x_1/r)\} + \dots,$$

and  $\kappa$  is the angular momentum operator

$$\kappa = \rho_3 \{(m, \sigma) + \hbar\}.$$

On multiplying (2.1) by  $(x, \sigma) \rho_1$  it becomes

$$\{(rp_r + i\rho_3\kappa) + c^{-1}(Er + e^2) \rho'_2\} \psi = \pm im_0cr\rho'_1\psi, \quad (2.2)$$

where

$$\rho'_1 = r^{-1}(x, \sigma) \rho_1,$$

and

$$\rho'_2 = r^{-1}(x, \sigma) \rho_2.$$

The operators,  $\rho'_1, \rho'_2, \rho_3$  form a new set of quaternions, which commute with every other operator in (2.2), and which can, therefore, be represented by Pauli's two-rowed matrices,

$$\begin{array}{ccc} 0 & 1 & 0 & -i & 1 & 0 \\ 1 & 0 & i & 0 & 0 & -1 \end{array}$$

The operator  $\kappa$  commutes with every other operator in (2.2) and hence can be replaced by its proper value  $\mu\hbar$ , where  $\mu = \pm 1, \pm 2, \dots$ , (the value zero

being excluded). Hence the equations for  $\psi_1$ ,  $\psi_2$ , the two components of  $\psi$ , in the Pauli representation, are

$$(r\partial/\partial r + 1 - \mu) \psi_1 + (\alpha + Er/\hbar c) \psi_2 = \mp (m_0 c/\hbar) r \psi_2,$$

and

$$(r\partial/\partial r + 1 + \mu) \psi_2 - (\alpha + Er/\hbar c) \psi_1 = \mp (m_0 c/\hbar) r \psi_1.$$

If  $a_1 r^\lambda$  and  $a_2 r^\lambda$  are the leading terms in the expansions of  $\psi_1$  and  $\psi_2$  in ascending powers of  $r$ , then

$$(\lambda + 1 - \mu) a_1 + \alpha a_2 = 0,$$

and

$$(\lambda + 1 + \mu) a_2 - \alpha a_1 = 0.$$

Hence

$$\lambda = -1 \pm (\mu^2 - \alpha^2)^{1/2}.$$

Now in order that the integral,

$$\int_a^b (\psi_1^2 + \psi_2^2) r^2 dr,$$

should converge as  $a \rightarrow 0$ , it is necessary that

$$2\lambda + 2 > -1, \quad \text{i.e.,} \quad \lambda > -3/2.$$

Since  $\mu$  cannot vanish, this condition allows only those solutions for which

$$\lambda = -1 + (\mu^2 - \alpha^2)^{1/2}.$$

In order that the integral should also converge as  $b \rightarrow \infty$ , it is necessary that  $E$  should satisfy the equation

$$\frac{\alpha E}{(m_0^2 c^4 - E^2)^{1/2}} = n_0 + (\mu^2 - \alpha^2)^{1/2}. \quad (2.3)$$

where  $n_0$  is a positive integer. Hence the difference between  $E$  and the proper energy  $m_0 c^2$  is given approximately by the equation

$$E - m_0 c^2 \doteq -(\alpha^2/2n^2) m_0 c^2, \quad (n = n_0 + |\mu|),$$

which is equivalent to Bohr's classical formula. The solutions with the specified values of  $\lambda$  and  $E$  are the wave functions for a hydrogen atom with serial quantum  $l$  and inner quantum number  $j$ , determined by the equations,

$$\begin{aligned} l &= \mu, & j &= \mu - \frac{1}{2}, & (\mu > 0), \\ l &= -\mu - 1, & j &= -\mu - \frac{1}{2}, & (\mu < 0). \end{aligned}$$

The first order equation possesses no other solutions which are quadratically integrable over the range,  $0 \leq r \rightarrow \infty$ .

3. *The Second Order Wave Equation.*

Since the first order equation has the form (2.1),

$$(F \pm im_0c) \psi = 0,$$

it is natural to take as the second order equation that which is obtained by removing the ambiguous sign of  $im_0c$ , i.e.,

$$(F \mp im_0c) (F \pm im_0c) \psi = 0,$$

or

$$F^2 \psi = -m_0^2 c^2 \psi.$$

Now

$$F^2 = (p_1^2 + p_2^2 + p_3^2) - c^{-2} (E + e^2/r)^2 + i (\hbar e^2/cr^3) \rho'_3. \quad (3.1)$$

This equation may be considerably simplified by introducing the operator

$$\gamma = -\rho_3 \kappa + i \hbar \rho'_3. \quad (3.2)$$

Since  $\rho'_3$  anti-commutes with  $\kappa$ , it follows that

$$\gamma^2 = \kappa^2 - \alpha^2 \hbar^2. \quad (3.3)$$

Now

$$\kappa^2 = (m_1^2 + m_2^2 + m_3^2) - \hbar (m, \sigma) + \hbar^2.$$

Hence

$$\gamma^2 + \hbar \gamma = (m_1^2 + m_2^2 + m_3^2) + i \alpha \hbar^2 \rho'_3 - \alpha^2 \hbar^2 \quad (3.4)$$

But

$$p_1^2 + p_2^2 + p_3^2 = p_r^2 + (m_1^2 + m_2^2 + m_3^2) r^{-2}.$$

Therefore

$$F^2 = p_r^2 + \frac{\gamma^2 + \hbar \gamma}{r^2} - \frac{E^2}{c^2} - \frac{2Ee^2}{c^2 r},$$

and the second order wave equation is

$$\frac{\partial^2 \psi}{\partial r^2} + \frac{2\partial \psi}{r \partial r} + \left\{ \frac{E^2 - m_0^2 c^4}{\hbar^2 c^2} + \frac{2\alpha^2 E}{c^2 r} - \frac{\gamma^2 + \hbar \gamma}{\hbar^2 r^2} \right\} \psi = 0. \quad (3.5)$$

$\gamma$  and  $\kappa$  commute, and their simultaneous proper values are of the form  $g\hbar$  and  $\mu\hbar$ , where

$$g = \pm (\mu^2 - \alpha^2)^{\frac{1}{2}}$$

and

$$\mu = \pm 1, \pm 2, \dots \quad (\mu \neq 0, i).$$

Hence, on replacing  $\gamma$  by its proper value, (3.5) reduces to

$$\frac{\partial^2 \psi}{\partial r^2} + \frac{2\partial \psi}{r \partial r} + \left\{ \frac{E^2 - m_0^2 c^4}{\hbar^2 c^2} + \frac{2\alpha^2 E}{c^2 r} - \frac{g(g+1)}{r^2} \right\} \psi = 0. \quad (3.6)$$

This equation is of the same form as Schrödinger's equation for the non-relativistic hydrogen atom, the theory of which is well known.

In view of the fact that the main object of Dirac's investigation was to construct a first order wave equation, linear alike in the energy  $E$  and the momentum operators  $p_1, p_2, p_3$ , it may appear a retrograde step to replace the first order equation by a second order equation. But the first order equation possesses no solutions other than those describing the orthodox stationary states of the hydrogen atom, as was stated at the end of the last section. Hence the first order equation is clearly incompetent to furnish a theory of the neutron, and it must be modified in some way if the neutron is to be fitted into the quantum theory as a kind of collapsed hydrogen atom. Once the need for some modification is admitted, it will be recognized that the construction of the second order equation is probably the simplest and most natural modification possible.

This second order equation

$$(F^2 + m_0^2 c^2) \psi = 0,$$

possesses, of course, all the solutions of the first order equation

$$(F \pm im_0 c) \psi = 0$$

These are the "H"-solutions, referred to in the Introduction. The second order equation also possesses another set of solutions, quadratically integrable over the range  $0 \leq r \rightarrow \infty$ . These are the "N"-solutions, which are here regarded as the wave functions of a neutron.

#### 4. The Stationary States of the Neutron.

It will now be shown that the energy levels of the "N"-solutions form a continuous spectrum, extending over the half-open interval,

$$-m_0 c^2 \leq E < m_0 c^2.$$

The solution for the ground level,  $E = -m_0 c^2$ , requires a separate discussion: the other solutions can be easily constructed as follows:—

If  $-m_0 c^2 < E < m_0 c^2$ , the substitutions,

$$\psi = r^{-1} W, \quad z = 2 \{ (m_0^2 c^4 - E^2) / (\hbar^2 c^2) \}^{\frac{1}{2}} r,$$

reduce the wave equation (3.6) to its canonical form,

$$\frac{d^2 W}{dz^2} + \left\{ -\frac{1}{4} + \frac{k}{z} + \frac{\frac{1}{4} - m^2}{z^2} \right\} W = 0, \quad (4.1)$$

where

$$k = \alpha E / (m_0^2 c^4 - E^2)^{\frac{1}{2}}, \quad m = |g + \frac{1}{2}|.$$

It is required to determine the solutions of this equation for which the integral

$$\int_a^b \psi^2 r^2 dr \quad \text{or} \quad \int_a^b W^2 dr, \quad (4.2)$$

converges as  $a \rightarrow 0$  and  $b \rightarrow \infty$ .

Two independent solutions of (4.1) are provided by Whittaker's functions,  $W_{k,m}(z)$  and  $W_{-k,m}(-z)$ , whose asymptotic behaviour as  $r \rightarrow \infty$  is specified by the relations,

$$W_{k,m}(z) \sim e^{-iz} z^k \{1 + O(z^{-1})\},$$

and

$$W_{-k,m}(-z) \sim e^{iz} (-z)^k \{1 + O(z^{-1})\}.$$

Hence, for the integrals (4.2) to converge as  $b \rightarrow \infty$ , the solution  $W$  must be a numerical multiple of  $W_{k,m}(z)$ .

As  $z \rightarrow 0$ ,  $W_{k,m}(z)$  is of order  $z^{k-m}$ , unless  $(k-m+\frac{1}{2})$  is equal to a positive integer  $n'$ , in which case  $W_{k,m}(z)$  is of order  $z^{k+m}$ . Hence the integral (4.2) will converge as  $a \rightarrow 0$  only in the following cases:—

(H)

$$k-m+\frac{1}{2} = n',$$

i.e.,

$$\frac{\alpha E}{(m_0^2 c^4 - E^2)^{\frac{1}{2}}} = n' + |g + \frac{1}{2}| - \frac{1}{2}; \quad (4.3)$$

and (N)

$$1 - 2m > -1,$$

i.e.,

$$|g + \frac{1}{2}| = m < 1.$$

Since  $g = \pm (\mu^2 - \alpha^2)^{\frac{1}{2}}$ , this inequality can be satisfied only if  $\mu^2 = 1$  and  $g = -(1 - \alpha^2)^{\frac{1}{2}}$ .

In case "H," the solutions yield the accepted wave functions of the hydrogen atom. As  $r \rightarrow 0$ , these functions are of order  $z^{m-\frac{1}{2}}$ . They therefore tend to zero as  $r \rightarrow 0$ , unless  $g = -(1 - \alpha^2)^{\frac{1}{2}}$  in which case they present a mild infinity of order  $r^{(1-\alpha^2)^{\frac{1}{2}}-1}$ , i.e., of a lower order than  $r^{-\frac{1}{2}}$ . Equation (4.3) for the energy levels is equivalent to equation (2.3) obtained from the first order equation.

In case "N," the solutions are quite distinct in character from the "H"-solutions. As  $r \rightarrow 0$ , the corresponding wave functions,  $\psi$ , become infinite of order  $r^{-(1-\alpha^2)^{\frac{1}{2}}}$ , but  $\psi^2$ , and even  $r^{-1} \psi^2$ , remains integrable over the range  $(0, \infty)$ . In this case the only limitations on  $E$  are the inequalities,

$$-m_0 c^2 < E < m_0 c^2,$$

so that the energy spectrum is continuous.



The value of the serial quantum number  $l$  for these solutions can be obtained by the following argument. It follows from equation (3.4) that, neglecting terms of order  $\alpha$  the operator  $(m_1^2 + m_2^2 + m_3^2)$  is approximately equal to  $(\gamma^2 + \hbar\gamma)$ . Hence, approximately,

$$l(l+1)\hbar^2 \doteq g(g+1)\hbar^2$$

and

$$l + \frac{1}{2} \doteq |g + \frac{1}{2}| = (1 - \alpha^2)^{\frac{1}{2}} - \frac{1}{2}.$$

Now  $l$  is a positive integer or zero. Therefore the accurate value of  $l$  is zero. Also, since  $\mu^2 = 1$  the inner quantum number  $j$  is  $|\mu| - \frac{1}{2}$  or  $\frac{1}{2}$ .

If  $E^2 > m_0^2 c^4$  the asymptotic value of the solutions of the wave equation (3.6) is of the form

$$\exp \{ \pm i (pr + q \log r) \}, \quad (p \text{ and } q \text{ real}).$$

Hence no solution of the equation can render the integrals (4.2) convergent as  $b \rightarrow \infty$ . Therefore, if  $E^2 > m_0^2 c^4$  there are no "closed" states. (The solutions in this case represent electrons fired at a stationary proton.)

### 5. The Normal State of the Neutron.

It remains to consider the solutions when  $E = \pm mc^2$ . In this case the wave equation (3.6) can be reduced to a standard form by the substitutions,

$$\psi = r^{-\frac{1}{2}} w, \quad x = r^{\frac{1}{2}},$$

which yield the equation,

$$\frac{d^2 w}{dx^2} + \frac{dw}{x dx} + \left\{ \frac{8\alpha^2 E}{e^2} - \frac{(2g+1)^2}{x^2} \right\} w = 0. \quad (5.1)$$

Two independent solutions of this equation are the Bessel functions

$$J_\nu \{ (8\alpha^2 E / e^2)^{\frac{1}{2}} x \},$$

of orders

$$\nu = \pm |2g + 1|.$$

If  $E = m_0 c^2$ , the arguments of the Bessel functions are real. Now, as  $z \rightarrow \infty$ ,  $J_\nu(z) \sim (2/\pi z)^{\frac{1}{2}} \cos(z - \frac{1}{2}\nu\pi - \frac{1}{4}\pi)$ . Hence, the asymptotic value of the wave function  $\psi$  is of the form

$$\psi \sim r^{-\frac{1}{2}} \cos(pr^{\frac{1}{2}} + q),$$

where  $p$  and  $q$  are real. Therefore the integral (4.2) cannot converge as  $b \rightarrow \infty$ . Hence  $E = m_0^2 c^2$  is *not* one of the energy levels allowed by the second order wave equation.

If, however,  $E = -m_0c^2$  the arguments of the Bessel functions are wholly imaginary. It is convenient to take as a new variable

$$z = (8\alpha^2 m_0 c^2 r / e^2)^{\frac{1}{2}} = (8r/a_0)^{\frac{1}{2}},$$

where  $a_0$  is the radius of the first Bohr orbit of hydrogen.

Equation (5.1) then becomes

$$\frac{d^2 w}{dz^2} + \frac{dw}{zdz} - \left\{1 + \frac{\nu^2}{z^2}\right\} w = 0,$$

where

$$\nu = |2g + 1|.$$

Two independent solutions of this equation are the Bessel functions  $I_\nu(z)$  and  $I_{-\nu}(z)$  defined by the series,

$$I_\nu(z) = \sum_{s=0}^{\infty} \frac{(\frac{1}{2}z)^{\nu+2s}}{s! \Gamma(\nu+s+1)}. \quad (5.2)$$

As  $z \rightarrow \infty$ , both of these functions have the asymptotic value  $(2\pi z)^{-\frac{1}{2}} e^z$ , whence the corresponding solutions of the wave equation (3.6) must also become infinite as  $r \rightarrow \infty$ . Hence the only combination of the Bessel functions which can lead to an admissible solution,  $\psi$ , is the difference,  $I_\nu(z) - I_{-\nu}(z)$ , which is of order  $z^{-\frac{1}{2}} e^{-z}$  as  $z \rightarrow \infty$ .

The standard function of this type is Macdonald's function

$$K_\nu(z) = \frac{1}{2}\pi \operatorname{cosec}(\pi\nu) \{I_{-\nu}(z) - I_\nu(z)\}, \quad (5.3)$$

which can also be expressed as the definite integral,

$$K_\nu(z) = \int_0^\infty e^{-z \cosh t} \cosh \nu t \, dt.$$

The corresponding solution of the wave equation (3.6) is

$$\psi(r) = A r^{-\frac{1}{2}} K_\nu \{(8r/a_0)^{\frac{1}{2}}\}. \quad (5.4)$$

( $A$  being a numerical constant). Hence the integral (4.2) will converge as  $b \rightarrow \infty$ .

As  $z \rightarrow 0$ ,

$$K_\nu(z) = O(z^{-\nu}),$$

whence,

$$\psi(r) = O(r^\lambda),$$

where

$$\begin{aligned} \lambda &= -\frac{1}{2}\nu - \frac{1}{2} = -|g + \frac{1}{2}| - \frac{1}{2} \\ &= -(\mu^2 - \alpha^2)^{\frac{1}{2}} - 1, \quad \text{if } g > 0, \end{aligned}$$

or

$$-(\mu^2 - \alpha^2)^{\frac{1}{2}}, \quad \text{if } g < 0.$$

Hence the integral (4.2) will converge as  $a \rightarrow 0$ , if and only if

$$g < 0, \quad \mu^2 = 1, \quad \nu = 2(1 - \alpha^2)^{\frac{1}{2}} - 1, \quad \text{and} \quad \lambda = -(1 - \alpha^2)^{\frac{1}{2}}.$$

As before, these values imply that the serial and inner quantum numbers of the normal state of the neutron are  $l = 0$  and  $j = \frac{1}{2}$ . The solution just obtained corresponds to a value of energy  $E$  equal to  $-m_0c^2$ . Hence the energy of the normal state of the neutron is

$$E_N = -m_0c^2.$$

It follows from (4.3) that the energy,  $E_H$ , of the normal state of the hydrogen atom, is given by

$$\alpha E_H = (m_0^2c^4 - E_H^2)^{\frac{1}{2}}$$

i.e.,

$$E_H = m_0c^2(1 + \alpha^2)^{-\frac{1}{2}}.$$

Hence the binding energy of the neutron is

$$E_N - E_H \doteq 2m_0c^2,$$

or about one million electron volts.

## 6. The Dimensions of the Neutron in its Normal State.

The spread of the electric charge in a neutron in its normal state can be estimated from the average values of  $r$  and  $r^{-1}$ , which can be calculated exactly as follows.

The numerical constant,  $A$ , in the expression for the wave function,

$$\psi(r) = Ar^{-K} \{ (8r/a_0)^{\frac{1}{2}} \},$$

must be chosen so that

$$4\pi \int_0^\infty \psi^2(r) r^2 dr = 1.$$

The average values of  $r$  and  $r^{-1}$  are then given by the integrals,

$$\bar{r} = 4\pi \int_0^\infty \psi^2(r) r^3 dr,$$

and

$$\bar{r}^{-1} = 4\pi \int_0^\infty \psi^2(r) r dr.$$

In terms of the dimensionless variable,  $z = (8r/a_0)^{1/2}$ , these three equations become

$$\begin{aligned}\bar{r}^{-1} &= 4\pi A^2 (2a_0/8) \int_0^\infty z K_r^2(z) dz, \\ 1 &= 4\pi A^2 (2a_0^2/8^2) \int_0^\infty z^3 K_r^2(z) dz, \\ \bar{r} &= 4\pi A^2 (2a_0^3/8^3) \int_0^\infty z^5 K_r^2(z) dz.\end{aligned}$$

The three integrals of the type,  $\int_0^\infty z^r K_r^2(z) dz$ , are evaluated in Appendix I.

On inserting their values it is found that

$$\begin{aligned}\bar{r}^{-1} &= (12/a_0) \cdot (1 - v^2)^{-1} \doteq 6/(\alpha^2 a_0) \\ A^2 &= \frac{1}{4\pi} \cdot \frac{8^2}{2a_0^2} \cdot \frac{3}{(1 - v^2)} \cdot \frac{\sin \pi v}{\pi v} \cdot \frac{12}{\pi a_0^2} \\ \bar{r} &= \frac{1}{10} a_0 (4 - v^2) \doteq \frac{1}{10} a_0\end{aligned} \quad (6.1)$$

The mean value of the electron-proton distance,  $\bar{r} \doteq 3a_0/10 \doteq 1.6 \times 10^{-9}$  cm. is decided less than the corresponding distance,  $\frac{1}{2} a_0$ , for the hydrogen atom in its normal state; but a much better idea of the effective size of the neutron is afforded by the mean value of the reciprocal of  $r^{-1}$ .

The value of  $\bar{r}^{-1}$  implies that the average electrostatic potential energy of the neutron in its normal state,  $-e^2/\bar{r}$ , is the same as if the electron were at a fixed distance  $\frac{1}{3}\alpha^2 a_0$  from the proton. Since  $\alpha \doteq 1/137$ , this distance is approximately  $10^{-5} a_0$ , or  $0.5 \times 10^{-13}$  cm. This distance is of the order of magnitude of the estimated radius of the neutron.

### 7. The External Field of the Neutron in its Normal State.

The average value of the electrostatic potential energy of a unit positive charge at a distance  $R$  from the nucleus of a neutron is equal to the potential of the nucleus,  $e/R$ , plus the potential due to a negative charge distributed with volume density  $\rho(r) = -e\psi^2(r)$ . This potential is the "external field" of the neutron. Its value for large values of  $R$  can be calculated from the asymptotic value of the wave function.

Since

$$K_r(z) \sim (\pi/2z)^{1/2} e^{-z}$$

it follows that

$$\psi(r) \sim A (4\pi/a_0)^{1/2} z^{-3/2} e^{-z},$$

and

$$\rho(r) \sim -48ea_0^{-3} z^{-3} e^{-2z}$$

using the approximate value of  $A$  given in equation (6.1). Hence the total charge outside a sphere of radius  $r$ , with centre at the nucleus, is approximately

$$(3\pi/8) ez^2 e^{-2z}.$$

The external field is therefore given, to a first approximation, by

$$V(r) = \frac{3\sqrt{2}}{16} \frac{e}{(a_0 r)^4} \exp \{ -2(8r/a_0)^{1/2} \}.$$

This expression is valid only when  $r$  is large compared with  $a_0$ .

The potential given by this formula decreases much less rapidly, as  $r \rightarrow \infty$ , than the corresponding potential assumed by Massey,\* viz.,

$$V(r) = e \left( \frac{1}{r} + \frac{Z}{a_0} \right) \exp(-2Zr/a_0), \quad (Z > 25,000),$$

but it is difficult to estimate the significance of the result deduced here, until it has been used to calculate the collision cross-section of a neutron.

## 8. Summary.

This paper develops a theory of the neutron on the basis of the second order wave equation for the hydrogen atom,

$$(F^2 + m_0^2 c^2) \psi = 0,$$

where  $F$  is Dirac's wave operator,

$$F \equiv (p, \sigma) \rho_1 + ic^{-1} (E + e^2/r) \rho_3.$$

It is shown that this equation possesses two types of solutions for which  $\int_0^\infty \psi^2 r^2 dz$  is finite—the type “H” which yields the accepted wave functions of the hydrogen atom, and the type “N” which is here identified with the wave functions of the neutron.

The following properties of the neutron are deduced from the form of the “N”-solutions:—

(1) The energy spectrum of the neutron covers the interval,

$$-m_0 c^2 \leq E < +m_0 c^2,$$

\* ‘Proc. Roy. Soc.,’ A, vol. 138, p. 460 (1932).

the energy of the normal state being  $-m_0c^2$ , and its binding energy, relative to the normal H-atom being, approximately,  $2m_0c^2$  or  $10^6$  electron volts.

(2) For every energy level of the neutron,  $l = 0$  and  $j = \frac{1}{2}$ . Hence the normal state of the hydrogen is really a metastable state from which no transitions can occur to the neutron states.

(3) The average value of  $r$ , the electron-proton distance in the normal neutron, is  $\bar{r} = 3a_0/10$ , where  $a_0$  is the radius of the first Bohr orbit of hydrogen. The average value of  $r^{-1}$  is  $\bar{r}^{-1} = 6/\alpha^2 a_0$ , where  $\alpha$  is the fine structure constant ( $\alpha \doteq 1/137$ ).

(4) The external field of the normal neutron is

$$V(r) = \frac{3\sqrt{2}}{16} \frac{e}{(a_0 r)^4} \exp \{-2(8r/a_0)^4\}.$$

#### APPENDIX I.

##### *Evaluation of the Integrals of the Form $\int_0^\infty z^\nu K_\nu^2(z) dz$ .*

The integrals required in § 6 are not explicitly given in Watson's "Treatise on Bessel Functions," but they can be easily evaluated by means of Lommel's integrals and Schafheitlin's reduction formula, given in §§ 5.11 and 5.14 of this book.

Since  $K \equiv K_\nu(z)$  satisfies the equation

$$z^2 K'' + zK' - (z^2 + \nu^2)K = 0,$$

it follows that

$$\frac{d}{dz} \{z^2 K'^2 - (z^2 + \nu^2) K^2\} = -2zK^2.$$

Hence

$$\int_0^\infty zK^2 dz = \frac{1}{2} [z^2 K'^2 - (z^2 + \nu^2) K^2].$$

Now the powers of  $z$  which appear in the series for  $K$ , (5.2) and (5.3) are

$$-\nu + 2s \quad \text{and} \quad \nu + 2s,$$

where  $s = 0, 1, 2, \dots$ . Hence the series for  $K^2$  and  $z^2 K'^2$  contain terms which are independent of  $z$ , and these terms determine the value of the integral, all terms in  $z^{-2\nu}$  cancelling, and all other terms tending to zero with  $z$ , since  $\nu \doteq 1 - \alpha^2$ . Therefore

$$\int_0^\infty zK^2 dz = \frac{\pi\nu}{2 \sin \pi\nu}.$$

The reduction formula for the second integral follows from the identity

$$\begin{aligned} \frac{d}{dz} \{z^3 K' K - z^2 K^2 - \frac{1}{2} z^2 [z^2 K'^2 - (z^2 + \nu^2) K^2]\} \\ \equiv 3z^3 K^2 + 2(\nu^2 - 1) z K^2. \end{aligned}$$

Now the expression differentiated is of order  $z^{2-2\nu}$  as  $z \rightarrow 0$ . Hence

$$\int_0^\infty z^3 K^2 dz = \frac{2}{3} (1 - \nu^2) \int_0^\infty z K^2 dz.$$

Similarly, it follows from the relation

$$\frac{d}{dz} \{z^4 [z^2 K'^2 - (z^2 + \nu^2) K^2] - 4z^5 K'K + 8z^4 K^2\}$$

that

$$\int_0^\infty z^5 K^2 dz = \frac{4}{5} (4 - \nu^2) \int_0^\infty z^3 K^2 dz.$$

## APPENDIX II.

*The "N"-solutions of Schrödinger's Wave Equation for the Hydrogen Atom.*

It is interesting to note that Schrödinger's non-relativistic wave equation for the hydrogen atom,

$$\frac{d^2\psi}{dr^2} + \frac{2d\psi}{rdr} + \left\{ \frac{2m_0E}{\hbar^2} + \frac{2}{a_0r} - \frac{l(l+1)}{r^2} \right\} \psi = 0,$$

possesses a second set of solutions, quadratically integrable over all space, in addition to its "H"-solutions, which give the non-relativistic wave functions of H.

It is easily seen that such solutions cannot exist if  $E > 0$ .

If  $E < 0$ , the transformation,

$$\psi = r^{-1} W, \quad z = (-8m_0E)^{\frac{1}{2}} r/\hbar,$$

reduces the wave equation to its canonical form,

$$\frac{d^2W}{dz^2} + \left\{ -\frac{1}{4} + \frac{k}{z} + \frac{\frac{1}{4} - m^2}{z^2} \right\} W = 0,$$

where

$$k = \alpha (-m_0c^2/2E)^{\frac{1}{2}}, \quad m = l + \frac{1}{2}.$$

Hence, as in § 4, it follows that admissible solutions exist only in the following cases :—

$$(H) \quad k - m + \frac{1}{2} = n' = 1, 2, 3, \dots$$

i.e.,

$$E/m_0c^2 = -\alpha^2/2(n' + l)^2;$$

and (N)

$$1 - 2m > -1,$$

i.e.,

$$l = 0.$$

In case "H," the solutions yield the accepted wave functions of the non-relativistic hydrogen atom. These solutions all remain finite as  $r \rightarrow 0$ . In case "N," the solutions are of the form,

$$\psi(r) = Az^{-1} W_{k, \frac{1}{2}}(z).$$

As  $r \rightarrow 0$ , they are of order  $r^{-1}$ . These solutions appear to represent a non-relativistic neutron, with serial quantum number  $l = 0$ , and with energy  $E$  lying in the continuous spectrum,  $E < 0$ .

If  $E = 0$ , the solutions of the wave equation are of the form

$$r^{-\frac{1}{2}} J_\nu \{(8r/a_0)^{\frac{1}{2}}\} \quad \text{and} \quad r^{-\frac{1}{2}} Y_\nu \{(8r/a_0)^{\frac{1}{2}}\},$$

where  $\nu = 2l + 1$ , and there are therefore no admissible integrals.

---



*Abstract*

545 . 21 541 . 132 . 3 : 547 . 466

*The Principle of Formaldehyde, Alcohol, and Acetone Titrations. With a Discussion of the Proof and Implication of the Zwitterion Conception.*

By GEORGE MAXWELL RICHARDSON

(From the Biochemistry Department, Imperial College of Science and Technology  
London, S.W.7)

(Communicated by C. R. Harington, F.R.S.—Received January 15, 1934.)

The principles upon which the formaldehyde, alcohol, and acetone titrations of amino acids depend have been discussed in the light of the zwitterion hypothesis and the evidence for this hypothesis has been summarized.

The Sørensen and the Foreman (Willstätter) titrations estimate amino groups, the Linderstrøm-Lang titration carboxylic groups, contrary to accepted convention. In the particular case of protein investigations, however, the old comparison of amino-N (Van Slyke) with Willstätter titrations to demonstrate equivalent liberation of acid and amino groups by enzymic hydrolysis remains valid, by a fortunate coincidence resulting from the modified titration procedure conventional in these studies.

Useful new applications of the general principles have been suggested.

*The full paper appears in 'Proc. Roy. Soc.,' B, vol. 115, p. 121.*

*Abstract*

543 . 846 . 062

*Critique on the Biological Estimation of Amino Nitrogen.*

By GEORGE MAXWELL RICHARDSON

(From the Biochemistry Department, Imperial College of Science and Technology,  
London, S.W.7)

(Communicated by C. R. Harington, F.R.S.—Received January 15, 1934.)

Comparison of titrimetric and Van Slyke methods of estimating amino-N in biological extracts has been made. Eight modifications of the formaldehyde and alcohol titration methods are included in an examination of plant leaf extracts.

All alcohol (or acetone) methods for amino-N are discarded for biological extracts owing to lack of specificity. Magnesium, which cannot be readily removed, and various weak acids, are titrated as amino-N by such methods.

Existing literature on the estimation of amino-N is summarized and the general usefulness and accuracy of the methods discussed. For coloured biological extracts, the Van Slyke method is preferred, but the formaldehyde (phenolphthalein) titration has advantages for colourless extracts. The implications of the zwitterion conception on the interpretation of titrimetric methods is stressed.

The effect of  $p_{\text{H}}$  on the efficiency of norite and lead acetate as decolourants of leaf extracts has been investigated. Lead acetate at  $p_{\text{H}}$  6 is adopted in preference to norite at  $p_{\text{H}}$  1. The colourstuff of the extract is found to migrate as an ampholyte during electro-dialysis.

*The full paper appears in 'Proc. Roy. Soc.,' B, vol. 115, p. 142.*

### *Abstract*

545. 21

## *The Differential Titration of Mixed Strong and Weak Acids. A Method for "Free HCl" in Gastric Contents, etc.*

By GEORGE MAXWELL RICHARDSON

(From the Biochemistry Department, Imperial College of Science and Technology, London, S.W.7)

(Communicated by C. R. Harington, F.R.S.—Received January 15, 1934.)

Current methods of estimating "free HCl" in acid mixtures are examined, reasons for preferring the Gunzberg method or a  $p_{\text{H}}$  2 titration being adduced.

These methods are superseded in accuracy and simplicity by a titration in 90% acetone solution to successive end-points at  $p_{\text{H}}$  3.4 and 4.0 using *p*-benzenesulphonic acid-azo- $\alpha$ -naphthylamine (described as naphthylamine orange) as indicator. The methods are compared on acid solutions of known composition.

*The full paper appears in 'Proc. Roy. Soc.,' B, vol. 115, p. 170.*

*Abstract*

545 . 37 : 547 . 466

*Electrometric Titration of Amino-acids in aqueous-alcoholic Solution.*

By ALBERT NEUBERGER

(From the Department of Pathological Chemistry, University College Hospital Medical School, London)

(Communicated by C. R. Harington, F.R.S.—Received February 8, 1934.)

(1) It has been shown that the acid-base theory of Brönsted can be made the basis of a single theoretical treatment of the constitution of ampholytes.

(2) A technique has been developed by which aqueous-alcoholic solutions of amino-acids may be titrated electrometrically against a standard aqueous calomel electrode with practically no interference from liquid junction or phase boundary potentials.

(3) Although the system used cannot be thermodynamically defined it is capable of showing clearly the variation in the behaviour of acidic and basic groups in relation to the dielectric constant of the solvent.

(4) The results obtained by this method give strong support to the zwitterionic theory of the constitution of amino-acids.

(5) The importance of electrometric titrations in aqueous alcohol for volumetric purposes and for the interpretation of dissociation constants has been pointed out and the basis of the existing volumetric methods of estimation of amino-acids has been discussed in the light of the facts revealed by the titration curves.

*The full paper appears in 'Proc. Roy. Soc.,' B, vol. 115, p. 180.*

---

*The Mechanism of Plastic Deformation of Crystals.*  
*Part I.—Theoretical.*

By G. I. TAYLOR, F.R.S., Royal Society Yarrow Professor.

(Received February 7, 1934.)

Experiments on the plastic deformation of single crystals, of metals and of rock salt have given results which differ in detail but possess certain common characteristics.

In general the deformation of a single crystal in tension or compression consists of a shear strain in which sheets of the crystal parallel to a crystal plane slip over one another, the direction of motion being some simple crystallographic axis. The measure of this strain, which will be represented by  $s$ , is the ratio of the relative lateral movement of two parallel planes of slip to the distance between them. Thus it is defined in the same way as the shear strain considered in the theory of elasticity.

The resistance to shear, which will be denoted by  $S$ , is defined as the component of shear stress in the direction of slip which must act parallel to the slip plane in order that plastic deformation may occur.

It has been found that when the results of tests on single crystals of a metal are analysed the stress-strain curve which represents  $S$  as a function of  $s$  is independent of the stress normal to the slip plane and of the components of shear stress perpendicular to the direction of slip. Thus the  $(S, s)$  curve is a unique curve which defines the strength of the single crystal at any stage of distortion.

When the  $(S, s)$  curves for single crystals of different metals are compared they are found to differ considerably in detail, but they all possess one general characteristic, a very small stress will produce a small plastic deformation,

With some crystals\* it has been found difficult to assign a definite stress at which plastic distortion begins, with others, such as rock salt and zinc, experimenters have found such a limit, but it is very small compared with the strength ultimately attained by the material. In some such crystals it has been found that the observed lower limit of strength depends very much on the degree of purity of the material, thus Schmid,† working with single crystals of

\* *E.g.*, aluminium.

† 'Z. Physik,' vol. 75, p. 538 (1932).

zinc, found that a decrease in the total amount of metallic impurity from 0.03% to 0.002% causes the stress at which the plastic deformation begins to decrease from 94 to 49 gm. per sq. mm.

Except for these differences in the early stages of distortion the  $(S, s)$  curves for many metallic single crystals are very similar. Some of them are shown in figs. 1, 2, and 3. Fig. 1 refers to aluminium. The data from which this curve

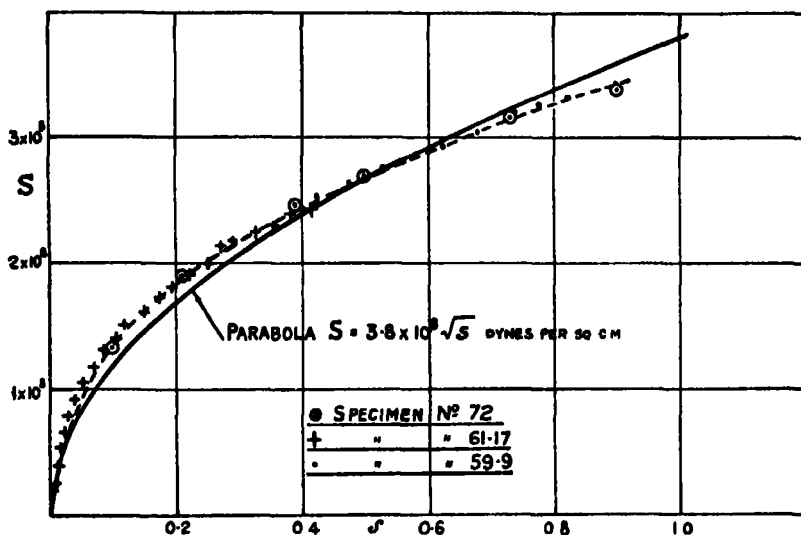


FIG. 1.—Aluminium crystals.

has been constructed were given in a previous paper.\* In fig. 1, however, the unit of stress has been changed from lbs. per square inch to dynes per square centimetre. Fig. 2 shows the result of tests on single crystals of copper and gold.† These tests were made by Dr. Elam and published in the form of curves giving  $S$  for various extensions of the specimens. The additional data necessary for calculating  $s$  were kindly furnished by Dr. Elam, and the results set forth in fig. 2 were calculated by means of the formulæ previously given.‡

Fig. 3 shows the results of tests on four single crystals of iron. The crystals in question were compressed and the geometry of their distortion analysed and published in a previous paper.§ The stress data required for the  $(S, s)$  curve had been recorded, but were not included in the published results.

\* Taylor, 'Proc. Roy. Soc.,' A, vol. 116, p. 51 (1927).

† Elam, 'Proc. Roy. Soc.,' A, vol. 112, p. 289 (1926).

‡ 'Proc. Roy. Soc.,' A, vol. 116, p. 47 (1927), equations (2), (4), (5), (6).

§ Taylor, 'Proc. Roy. Soc.,' A, vol. 112, p. 338 (1926).

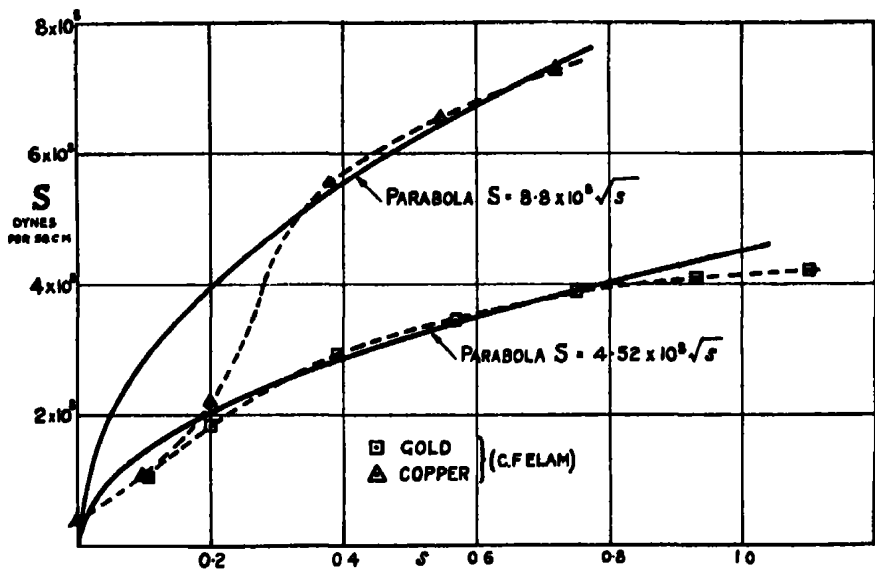


FIG. 2.—Copper and gold crystals.

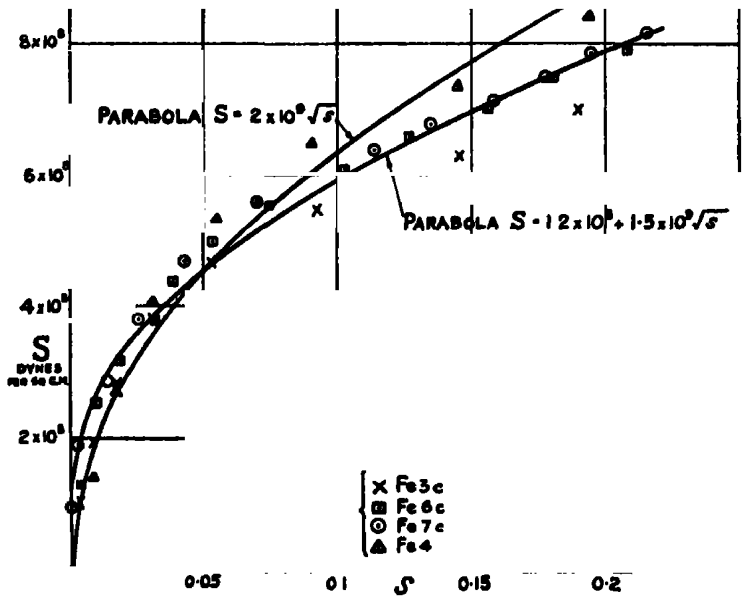


FIG. 3.—Iron crystals.

It will be seen that for iron and aluminium at any rate the curves resemble the parabolas.

Such parabolas have been drawn in figs. 1, 2, and 3 in order that they may be compared with the results of observation. For copper, fig. 2, the first three observed points do not lie on the parabola, while for iron, fig. 3, it seems that the experiments indicate a finite value (about equal  $1.2 \times 10^8$  dynes per sq. cm.) for the stress at which plastic distortion first occurs. The parabolas  $S = 1.2 \times 10^8 + 1.5 \times 10^9 \sqrt{s}$  and  $S = 2.0 \times 10^9 \sqrt{s}$  have been marked on the figure for comparison with the measured stresses.

### *Theories of Strain Hardening.*

Several attempts have been made to explain why plastic deformation increases the strength of metals (and rock salt), but they have mostly been of a qualitative nature. Such explanations are principally of three types. One type uses the observed fact that, when a single crystal breaks down under load, small portions of it are rotated into an orientation differing from that of the main body of the crystal. It is supposed that the perfect crystal can slip under the action of a very small stress, but that the portions of the crystal which are rotated relatively to the main body act as locking keys and hold up the slipping of the surrounding crystal, enabling the whole system to withstand a much greater load than the perfect crystal can support.

When the crystal contains *only* one set of planes of easy glide parallel to one crystallographic plane this theory might be capable, after some modification, of explaining a very rapid strengthening with increasing distortion. Where, however, the portions inside and outside a surface of misfit are supposed to be perfect crystals, but with crystallographic axes oriented in two different directions, easy gliding in the outer crystal would produce a definite strain (*i.e.*, alteration of shape) in the surface of misfit while easy gliding in the inner crystal would produce a different strain. Thus the only types of strain which are possible as a result of easy gliding in both the inner and the outer crystal are geometrically inconsistent at the boundary between them. If no other type of plastic strain is possible, then the inner crystal completely locks the portion of the outer crystal within which the planes of easy glide cut any portion of the inner crystal.

When the crystals are capable of slipping equally easily on several crystallographically similar planes, as, for instance, when the crystals have cubic symmetry, this explanation of hardening due to cold work is no longer applicable,

because, by an appropriate combination of easy glides, the inner crystal can be given exactly the same strain as the outer one. Under these conditions, therefore, there would be no geometrical inconsistency at the surface between them.

It is true that the stress necessary to cause any given kind of strain is likely to vary slightly according to whether the crystallographic axes are arranged in the most favourable or the most unfavourable orientation for the particular type of strain considered, but this effect is necessarily very small compared with the very large effect which the theory was designed to explain.

We may now turn to a second type of explanation. A perfect crystal is again supposed to be capable of withstanding only a very small stress, but every real crystal contains surfaces of misfit which form a mosaic structure or, as some crystallographers have described them, a system of lineages. These surfaces of misfit are supposed to hold up slipping to an increasing extent as the amount of distortion proceeds. This explanation, though at first sight promising, entails very great difficulty as soon as it is examined in detail.

The idea that the boundary can hold up slip in the interior of a perfect portion of the crystal might be interpreted in two ways. In one the surfaces of misfit may be regarded as being capable by themselves of carrying the whole applied stress, the intervening portions of perfect crystal structure being in a stress-free state. The strength would then be like that of a honey-comb, depending entirely on the boundaries which receive no assistance from the contents of each cell. Even if it were possible to explain the existence of any given strength on this principle it would be extremely difficult to explain why this increases with the amount of plastic strain.

In the other interpretation any thin sheet bounded by two parallel slip planes must move as a whole. Each perfect portion of the crystal can support any stress system for which there is no component of shear stress parallel to the slip planes, and the varying orientations of neighbouring portions which are crystallographically perfect might enable the whole system to support a large total stress in spite of the fact that each portion was excessively weak to one type of shear stress. This interpretation would make theory identical with the first type already discussed. It breaks down if the crystal has several crystallographically similar possible slip planes.

The third type of explanation is the inverse of the first two. A perfect crystal is supposed to be capable of withstanding a very large stress. The observed weakness of metal crystals is attributed to concentrations of stress due to internal surfaces of misfit or cracks, and the increasing strength with



increasing plastic strain is attributed to an increase in the number of faults or cracks.\* As the number of such faults increases the ratio of maximum stress in a region of stress concentration to mean stress in the material would be expected to decrease. The mean stress necessary to cause a given maximum stress in a region of stress concentration must, therefore, increase as the number of faults per unit volume of the material increases.

In the present paper an attempt will be made to present a theory of this kind in a quantitative form. It will be found that the idea that a system of faulting exists in the structure of an apparently perfect crystal arises naturally in the course of the work.

### *Significance of Direction of Slip in Crystals.*

Perhaps the most remarkable feature of the distortion phenomena of single crystals is that during the whole process of slipping parallel to one crystal plane they slip in a definite crystallographic direction. Regarding the matter from the atomic point of view the finite rigidity of the deformed crystal implies that each atom is situated in a position of stable equilibrium in the field of force due to its neighbours, and to intervening free electrons if there are any. Plastic strain must be due to stability interchanges in which atoms jump, owing to thermal or other agitation and to the applied stress, from one position of equilibrium to another. If these jumps are due to thermal agitation, they must be regarded as random occurrences though, since each jump alters the field of force in its neighbourhood, the probability that any given jump will occur may change very greatly when a jump occurs in a neighbouring position. From this point of view the whole macroscopic phenomenon of gliding must be regarded as the integrated effect of individual jumps, and since the direction of the resultant glide is parallel to a line of atoms in the crystal, it is difficult to avoid the conclusion that a great majority at any rate of the individual jumps are in this direction too.

If this is so, the length of the jump is likely to be equal to the spacing of atoms along this line, and the mechanism of slipping may be more like the simple shift shown in fig. 4 (c), in which the whole of the material on one side of a definite plane shifts through the length of one lattice cell, than has often been supposed. It is proposed, therefore, to examine how this ideal condition differs from what is observed in real materials. In the first place this ideal slipping would leave the material in the form of a perfect crystal and the

\* Taylor, 'Trans. Faraday Soc.,' vol. 24, p. 121 (1928).

strength would be unaltered by the distortion. In the second place to shift the whole of the upper row of atoms simultaneously over the lower row would necessitate the application of a stress comparable with, though no doubt less than, the elastic moduli of the material. Further, it would leave no room for an explanation of the large observed effect of temperature on plastic distortion.

*Slipping Over a Portion of the Slip Plane.*

It seems that the whole situation is completely changed when the slipping is considered to occur not simultaneously over all atoms in the slip plane but over a limited region, which is propagated from side to side of the crystal in a finite time. It has been observed by Joffe\* that the first sign of breakdown when a stressed crystal of rock salt is examined between nicol prisms is the appearance of a bright line, indicating distorted material, which is propagated along a slip line from side to side of the crystal. The stresses produced in the material by slipping over a portion of a plane are necessarily such as to give rise to increased stresses in the part of the plane near the edge of the region where slipping has already occurred, so that the propagation of slip is readily understandable and is analogous to the propagation of a crack.

If, in order to retain the explanation of the observed constancy of the direction of slip, we assume that the propagation of a line of slipping along a slip plane leaves a perfectly well ordered crystal arrangement in its wake, we obtain a system represented pictorially in fig. 4. In the diagram (a) represents the atoms in the lattice of a crystal block; (b) the condition of this block when a slip of one atomic spacing has been propagated from left to right into the middle; and (c) the block after the unit slip, or "dislocation" as we may call it, has passed through from left to right.

*Positive and Negative Dislocations.*

The configurations *a*, *b*, and *c* of fig. 4 represent the passage of a dislocation in which the atoms above the operative slip planes are compressed in the direction of slip while those below are expanded. It is clear that a similar dislocation can exist which is the mirror image in the slip planes of that shown in *b*, fig. 4. This configuration which is shown in fig. 4 (*e*) will be called a negative dislocation to distinguish it from fig. 4 (*b*) which will be called a positive dislocation. The passage of a positive dislocation across a crystal

\* "The Physics of Crystals," p. 47.

block from left to right produces the same effect as the passage of a negative one from right to left, namely, the shift of the upper part of the block through one atomic spacing to the right relative to the lower part. This is illustrated in fig. 4 by the sequence of configurations *d, e, f*.

It seems likely that the motion of a "unit dislocation" may be determined by the stress and the temperature. It is known that in the case of metals there is a recrystallization temperature, below the melting point, at which large irregularities in structure such as those which must exist at the boundaries of crystal grains with different crystal orientations can be propagated through

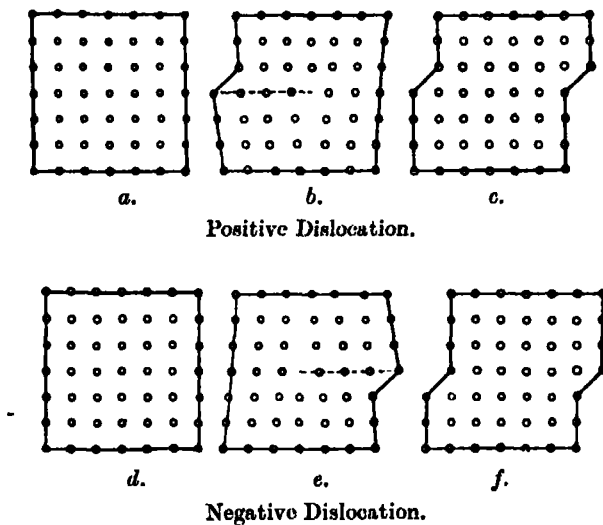


FIG. 4.—Positions of atoms during the passage of a dislocation.

the material. It may well happen therefore that there is a definite temperature probably lower than the recrystallization temperature, at which a unit dislocation of the type described above may be capable of free movement in either direction along the crystal plane. Let us call this hypothetical temperature  $T_D$ .

This picture seems adequate to account for the fact observed at high temperatures that slipping may occur at very low stresses in single crystals and that, after an indefinite amount of distortion, the crystal is still nearly in its original condition, except for change in shape; it is still a single crystal and still has very little strength. The distortion would then be accounted for by the propagation of centres of dislocation along the slip planes. These would enter the specimen at one side and leave it at the other as shown in fig. 4.

*Atomic Model of a Dislocation.*

Some insight into the manner in which a dislocation might be propagated along a slip plane may be gained by introducing a conception due to Dehlinger\* to explain recrystallization and described by him as "Verhakung" or "hooking." In theories of the equilibrium of crystal lattices an atom is supposed to place itself in a position of minimum potential energy in relation to its neighbours. Heat motions will agitate the atom so that it moves about in the neighbourhood of this position of minimum potential energy, but, until a certain temperature is reached, the chance that the atom will escape across the "potential barriers" which surround it is extremely small. In a perfect crystal structure one might, for instance, consider the potential along a line CD spaced midway between two regularly spaced lines of atoms  $A_0, A_1, A_2$  and  $B_0, B_1, B_2$ , fig. 5.

If the positions of minimum potential due to each of the lines considered separately occur at the points  $C_0, C_1, C_2$ , where  $C_0$  is midway between  $A_0$  and  $B_0$  atoms placed at these points would remain in equilibrium there and the equilibrium of a cubic structure might be partially explained in this way.

For the present purpose we may assume that the potential along CD due to either row can be represented approximately by

$$-A \cos (2\pi x/\lambda), \quad (1)$$

where  $\lambda$  is the lattice spacing, and  $x$  is measured from the position  $C_0$  midway between  $A_0$  and  $B_0$ .

The total potential due to both rows, namely,  $A_0, A_1, A_2, \dots$  and  $B_0, B_1, B_2, \dots$ , is then

$$\phi = -2A \cos (2\pi x/\lambda). \quad (2)$$

The height of the potential hill or barrier which must be climbed by any atom in order to escape altogether from its equilibrium position is then  $4A$ . This is represented graphically by the sine curve of amplitude  $2A$  which is shown in fig. 5.

In the neighbourhood of a positive dislocation the atoms in the slip plane above the centre of dislocation are compressed in the direction of slip while those in the plane below it are extended. At some distance from the centre the spacings along lines of atoms above and below the path of the centre of dislocation become regular and equal to the normal spacing for a perfect

\* 'Ann. Physik,' vol. 2, p. 749 (1929).

crystal, but if the number of atoms in a length  $L$  of a line of atoms which passes below the centre of a unit dislocation is  $N$ , then the number in the same length  $L$  on a line which passes above the centre is  $N + 1$ .

To calculate the distribution of potential on the line of atoms along which the actual centre of dislocation passes it would be necessary to solve the very difficult problem of finding the equilibrium positions of all the surrounding atoms. Though this is not yet possible the general nature of the potential

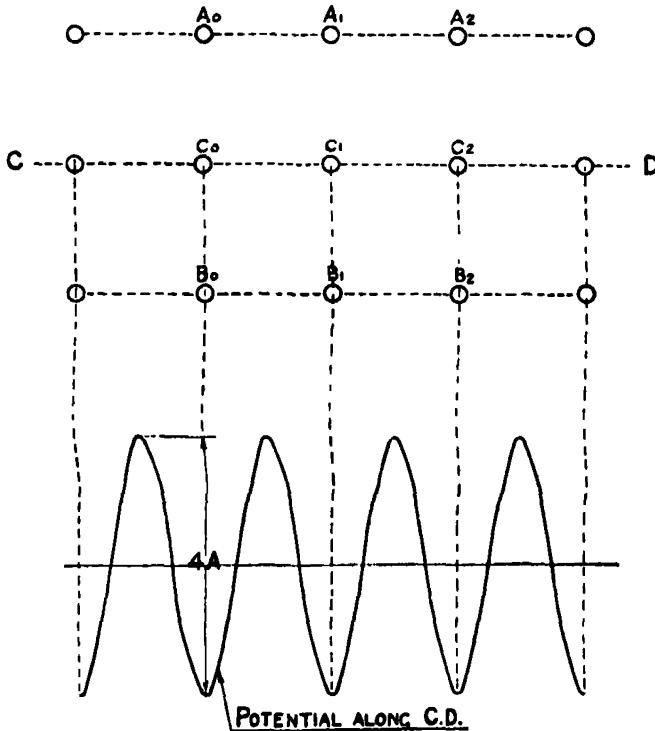


FIG. 5.—Arrangement of atoms in a perfect crystal structure.

distribution on the line along which the centre of the dislocation passes may perhaps be inferred from the results of calculation of potential due to an arbitrary distribution of atoms which has the essential characteristic that  $N + 1$  atoms in a line above the centre correspond with  $N$  atoms below it.\* The simplest arrangement of this kind is when  $N + 2$  atoms in the upper line are evenly spaced over a length  $(N + \frac{1}{2}) \lambda$  while the  $N + 1$  atoms in the lower line are evenly spaced over the same length. This arrangement is shown in

\* These numbers include the atoms at each end of the line so that  $N + 1$  of the compressed atomic spacings in the upper row correspond with  $N$  spacings in the extended lower row.

fig. 6 for  $N = 3$ . Outside this range we may suppose the two rows to have the spacing  $\lambda$  of the unstrained crystal. Making the same approximation as that for a regular lattice, the potential along the line CD midway between these two lines is

$$P = -A \cos 2\pi \frac{x}{\lambda} \left( \frac{N+1}{N+\frac{1}{2}} \right) - A \cos 2\pi \frac{x}{\lambda} \left( \frac{N}{N+\frac{1}{2}} \right). \quad (3)$$

The nature of this function is well known. It is shown graphically in fig. 6 for the case  $N = 3$ . In the length  $(N + \frac{1}{2}) \lambda$  it has  $N + 2$  minima including those at the two ends of the range. The heights of the "potential hills"

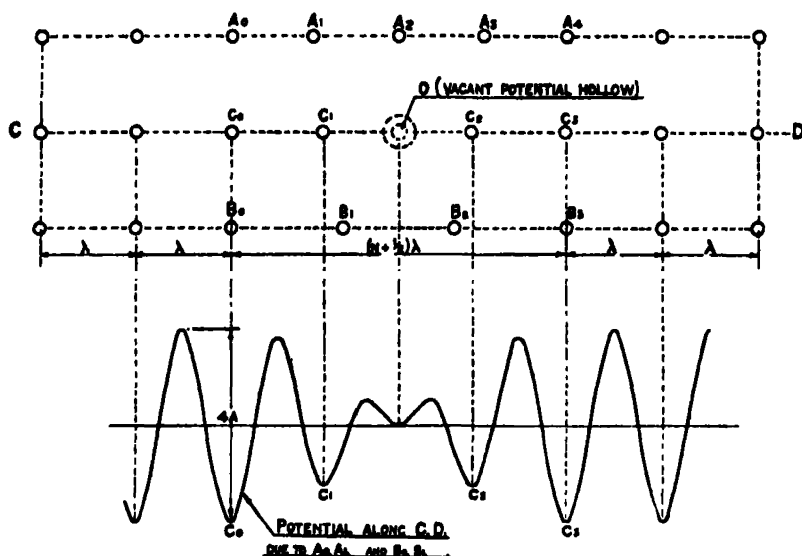


FIG. 6.—Conditions at dislocation when no external stress is applied.

and the depths of the "potential valleys" decrease from each end to the centre of the range. If  $N$  is an odd number, the shallowest depression on the potential curve is in the centre of the range. If we now imagine that each of the potential depressions except the central one is occupied by an atom we have a rough picture of a possible equilibrium distribution of atoms in the neighbourhood of a centre of dislocation. The actual centre of the dislocation may be regarded as being situated at the point O, fig. 6, in the middle of the central vacant potential hollow.

It has been pointed out that in a perfect crystal the atoms may be expected to remain within their potential barriers till a certain temperature is reached at which they can jump these barriers. At this temperature the atom  $C_0$ , fig. 6, in the disturbed lattice might be able to jump into the space occupied

by  $C_1$ . At lower temperatures an atom would not be capable of surmounting potential hills between  $C_0$  and  $C_1$ , but might still be able to jump the lower hills nearer the centre of the disturbance. The lowest temperature at which a jump can be made is that which will just permit either of the two central atoms to jump one of the lowest hills into the vacant central hollow O, fig. 6. This temperature we may call  $T_D$ . It may be noticed that even at the absolute zero of temperature some energy of agitation still persists in the crystal. If this is large enough to permit jumps to take place, then  $T_D = 0^\circ$  abs.

When one of the two central atoms has jumped into the central hollow, the potential hollow which is left vacant now becomes the centre of the dislocation and we must imagine that the neighbouring atoms readjust themselves by a continuous process (*i.e.*, without jumps) into new positions of equilibrium so that the potential hollow newly vacant becomes the shallowest one. This idea has already been expressed, without attempting to give it a definite atomic interpretation, in the statement made on p. 369 that a temperature  $T_D$  exists above which the centres of dislocation can move freely in either direction along the slip plane.

The foregoing picture is very incomplete because it takes no account of the mutual action of the atoms in the row along which the jumps occur, but the theory of the equilibrium of the atoms in *perfect* metallic crystals is so incomplete that it does not seem worth while to attempt any more complete discussion of equilibrium in a disturbed lattice.

### *Effect of Applying an External Shear Stress.*

We are now in a position to examine, in the light of the foregoing rough representation of atomic structure, the effect of applying an external shear stress. We may imagine that a shear stress applied in the direction indicated by the arrows in fig. 7 gives the row of atoms  $A_0, A_1, A_2, \dots$  a small displacement to the right relative to the row  $B_0, B_1, B_2, \dots$ . If this small displacement is  $\delta$ , the potential along the line CD is now

$$P = -A \cos 2\pi \frac{x - \delta}{\lambda} \left( \frac{N+1}{N+\frac{1}{2}} \right) - A \cos 2\pi \frac{x}{\lambda} \left( \frac{N}{N+\frac{1}{2}} \right). \quad (4)$$

It will readily be seen that the effect of this is to lower the height of the hill which the atoms to the right of the central hollow must jump in order to enter the central hollow. The hill which the atoms to the left of the centre must jump is correspondingly raised. The actual potential curve for  $N = 3$ ,

$\delta/\lambda = 0.1$  is shown in fig. 7. The positions of the atoms and of the vacant potential hollow or centre of the dislocation are also shown in fig. 7. Here height of the potential hill  $h_1$ , fig. 7, which the atom  $C_1$  to the left of the centre of dislocation has to jump in order to land in the central potential hollow is  $2.70 \text{ \AA}$ . The atom  $C_2$  to the right of the centre has only to jump a hill of height  $h_2 = 0.85 \text{ \AA}$ .

If the temperature is high enough to permit the jump  $h_2$  but not high enough to permit the jump  $h_1$ , the only jump which can occur is to the left into the centre of dislocation from the potential hollow on its right. The actual centre, *i.e.*, the vacant hollow, can therefore only move to the right. At higher

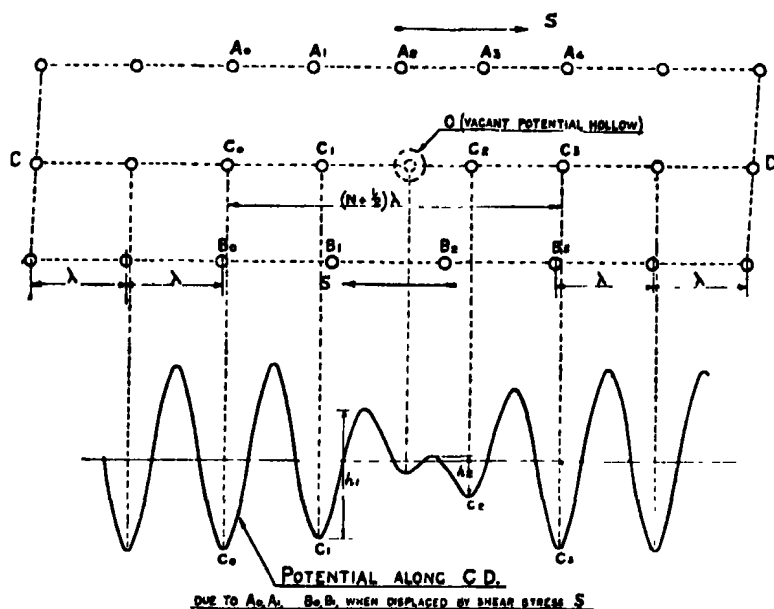


FIG. 7.—Conditions at dislocation in crystal subjected to shear stress.

temperatures, when the jumps  $h_2$  and  $h_1$  can both occur, the former will be the more frequent so that a positive centre of dislocation will necessarily migrate to the right under the action of a shear stress directed to the right. Similarly a negative dislocation will move to the left under the action of the same shear stress.

When the temperature is high enough to permit free movement of the centre of dislocation in the absence of applied shear stress the application of the smallest shear stress will cause the centre of dislocation to migrate along the slip plane. Thus plastic strain will occur when the smallest stress is applied. At lower temperatures a finite shear stress may be necessary in



order that the potential hill on one side of the centre of dislocation may be reduced in height to such an extent that the neighbouring atom can jump into the central hollow.

In the foregoing discussion an atomic mechanism is roughly represented which would give metallic crystals the following properties :—

- (1) At temperatures above  $T_D$  any stress, however small, applied parallel to the slip plane in the positive direction (see fig. 7) will cause positive centres of dislocation to travel in the positive direction and negative centres to travel in the negative direction ;
- (2) At temperatures below  $T_D$  the centres will not move till the shear stress attains some finite value.

In the work which follows these properties will be assumed and some possible consequences of their existence will be explored. It must be remembered that (1) is an accurate statement of what is to be expected of the atomic mechanism, but (2) is only a rough representation of some more complicated relationship involving temperature and rate of deformation which would result from a more complete consideration, from the statistical point of view, of the connection between temperature and the frequency of the stability interchanges, to which the motion of a dislocation is here ascribed.

#### *Distribution of Stress near a Unit Dislocation.*

The distribution of stress due to a unit dislocation cannot be calculated at points within a few lattice spacings of the centre of the disturbance, but at greater distances the theory of elasticity must apply. The general theory of dislocations in continuous isotropic elastic solids has been treated extensively by Volterra.\* In the neighbourhood of a centre of dislocation the stresses become very large so that it is necessary to imagine that the material is cut away round the actual centre. We may therefore suppose that our material is initially unstrained and that it contains a hole H, fig. 8. From some point A on the boundary a centre of dislocation passes along a straight line AB to the point B which is on the surface of the hole H.

If P is a particle just above AB and P' is just below it, they suffer a relative displacement  $\lambda$  during the passage of the centre of dislocation between them.

The appearance of a circular ring which has suffered this kind of strain is shown in fig. 8. The stress distribution round such a dislocation has been

\* 'Ann. Éc. Norm.,' vol. 24 (1907).

calculated by Timpe.\* It depends on the surface tractions applied at the surface of the hole  $H$  as well as on the amount of slip; but if the resultant force and couple due to these surface tractions are both zero the stresses due to the dislocation vary as  $r^{-1}$ ,  $r$  being the distance from the centre  $O$ , while the stresses due to the surface tractions vary as  $r^{-2}$  or higher inverse powers.

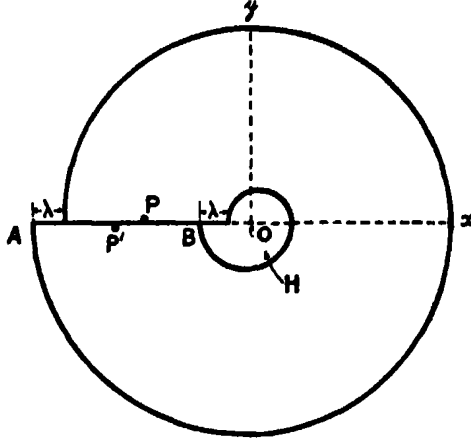


FIG. 8.—Elastic cylinder with dislocation.

At points far from the centre the components of stress due to a positive dislocation or slip of amount  $\lambda$  on the positive axis of  $x$  are

$$X_x = -Y_y = -\frac{\mu\lambda}{\pi} \frac{y}{x^2 + y^2}, \quad X_y = \frac{\mu\lambda}{\pi} \frac{x}{x^2 + y^2} \quad (5)$$

or in polar co-ordinates

$$\hat{r}\hat{r} = -\hat{\theta}\hat{\theta} = +\frac{\mu\lambda \sin \theta}{\pi r}, \quad \hat{r}\hat{\theta} = \frac{\mu\lambda \cos \theta}{\pi r}, \quad (6)$$

where  $\mu$  is the modulus of rigidity.

In the particular instance when no traction is exerted on the surface of a circular hole  $H$  of radius  $a$ , the complete expressions for the components are

$$\left. \begin{aligned} \hat{r}\hat{r} &= +\frac{\mu\lambda}{\pi r^3} \sin \theta (r^2 - a^2) \\ \hat{r}\hat{\theta} &= \frac{\mu\lambda \cos \theta}{\pi r^3} (r^2 - a^2) \\ \hat{\theta}\hat{\theta} &= -\frac{\mu\lambda}{\pi r^3} \sin \theta (r^2 + a^2) \end{aligned} \right\} \quad (7)$$

These are identical with (6) when  $a$  is small compared with  $r$ .

\* 'Göttingen Diss.,' Leipzig (1905).

It appears that whatever may be the state of stress at distances from the dislocation centre comparable with  $\lambda$  the stress at large distances is represented by (5). Thus, even though nothing is known about the exact conditions of equilibrium of atoms at a dislocation, the stress distribution at a distance of several lattice cells away is known, for it depends only on the amount of the slip. The application of the theory of dislocations in an elastic body to a crystal is, strictly speaking, only justifiable if the crystal is elastically isotropic, a condition which appears to be fulfilled for aluminium. In cubic crystals the error due to want of isotropy is likely to be small. It will be left out of consideration in the following pages.

*Mutual Action and Equilibrium of Centres of Dislocation.*

We may consider a single set of parallel planes as possible slip planes. If these are taken as perpendicular to the axis of  $y$ , the application of a positive shear stress will cause a positive centre to move to the right or a negative centre to the left. For the present we may suppose that the temperature is above  $T_D$  so that the dislocation moves, however small  $X_y$  may be. Since each centre of dislocation produces a distribution of shear stress in its neighbourhood, every centre must tend to cause all neighbouring centres to migrate along their slip planes.

Referring to equation (5) it will be seen that in the neighbourhood of a positive centre of dislocation,  $A$ ,  $X_y$  is positive when  $x$  is positive, if therefore there is another positive centre  $B$  whose  $x$  co-ordinate relative to  $A$  is positive,  $B$  will move to the right under the influence of  $A$ . That is, positive centres of dislocation repel one another. Similarly two negative centres repel one another. On the other hand, positive centres attract negative centres and *vice versa*. If there are only two centres in the whole field and they are both of the same sign, they will repel one another to an infinite distance (provided the temperature is above  $T_D$ ). Centres of opposite sign will attract one another till the line joining them is perpendicular to the slip plane, when they will rest in equilibrium under their mutual influence. This position is illustrated in fig. 9 (a).

If two centres of opposite sign lie on the same slip plane, they will attract one another, but, since the elastic equations cannot be applied when they come within a few lattice distances of one another, it is not possible to say whether or not they will neutralize one another and leave a perfect crystal structure.

We may now consider the effect of applying a positive shear stress  $S$  so that the upper face of a crystal block is subjected to a tangential stress acting towards the right. If only one dislocation is present, it will move across the block, to the right if it is positive and to the left if it is negative, however small  $S$  may be. In either case the result is the same, namely, a shift of the upper portion of the block relative to the lower part through a distance  $\lambda$ . Apart from this shift the properties of the block are unchanged by the passage of the dislocation.

Conditions are very different when two dislocations of opposite sign are present. Suppose, for instance, that owing to the application of a very small

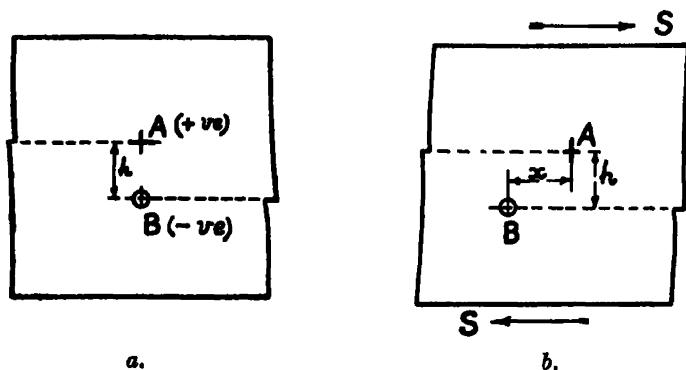


FIG. 9.—Positive and negative dislocations in equilibrium. (a) Under no external stress. (b) when shear stress is applied.

shearing force a positive dislocation starts moving along a slip plane from the left, and a negative dislocation starts to move from the right along another slip plane, distant  $h$  from the first. Even without application of the shear stress they would arrive under their mutual influence at the positions A, B, fig. 9 (a) where AB is perpendicular to the slip planes. On the other hand, in order to make them continue to move along their respective slip planes beyond AB, a shear stress must be applied. A shear stress  $S$  will cause the positive dislocation to move to the right and the negative one to the left till

$$S = \frac{\mu \lambda x}{\pi (x^2 + h^2)}, \quad (8)$$

where  $x$  is the projection on the slip plane of the distance between them. In this position each dislocation is in equilibrium.

The connection between  $S$  and  $x$  is shown in fig. 10. It will be seen that as  $x/h$  increases  $S\pi h/\mu\lambda$  increases till the maximum value 0.5 is reached at  $x = h$ . The maximum possible value of  $S$  consistent with equilibrium is therefore

$\frac{1}{2} \frac{\mu\lambda}{\pi h}$  or  $0.16 \mu\lambda/h$ . If  $S$  is less than this, the two centres of dislocation cannot escape from one another, whereas when  $S > \mu\lambda/2\pi h$  they pass one another and escape from their mutual attraction. Thus plastic distortion or slipping can only proceed, so far as the movement of these two centres of dislocation are concerned, when  $S > \mu\lambda/2\pi h$ . This is remarkable in view of the fact that the strength of the material is zero when only one centre of dislocation is present. Since  $S$  is inversely proportional to  $h$  it appears that slipping will

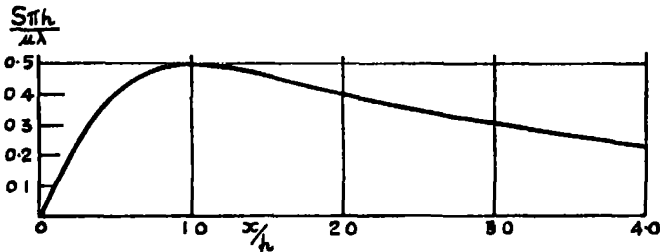


FIG. 10.

proceed on the slip planes which are farthest from planes of slip already containing centres of dislocation.

We may now consider possible equilibrium arrangements of centres of dislocation. It has been seen that positive and negative centres cannot exist on the same slip plane, for they would run together and neutralize one another. The simplest arrangements in which every centre is in equilibrium with its neighbours are those in which all the positive centres and all the negative

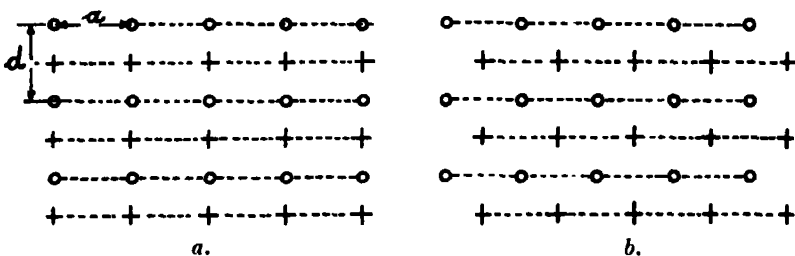


FIG. 11.—Equilibrium arrangements of dislocations, (a) stable, (b) unstable.

centres lie on interpenetrating rectangular nets. Two possible equilibrium configurations are shown as (a) and (b), fig. 11, where the positive centres are represented by crosses and the negative ones by round spots. The application of a positive shear stress to a crystal containing such a system would move all the positive centres to the right and all the negative ones to the left along the slip planes, represented by broken lines in fig. 11.

We shall calculate the displacement of the positive lattice relative to the negative when a given shear stress  $S$  is applied to the crystal. For this purpose we must first calculate the distribution of  $X_y$  due to the complete lattice of dislocations. The shear stress due to a row of positive dislocations regularly spaced at distance  $d$  apart in a line perpendicular to the slip planes is

$$\frac{\mu\lambda}{\pi} \sum_{n=-\infty}^{+\infty} \frac{x}{x^2 + (y + nd)^2} \quad (9)$$

where  $x$  is the distance of the point considered from the plane of the dislocations and  $y$  is its ordinate relative to an axis which passes through one of them. The sum of this series is known. It is

$$\frac{\mu\lambda}{d} \frac{\sinh(2\pi x/d)}{\cosh(2\pi x/d) - \cos(2\pi y/d)}; \quad (10)$$

when  $y = 0$  this becomes

$$\mu\lambda d^{-1} \coth(\pi x/d); \quad (11)$$

when  $y = \frac{1}{2}d$  it is

$$\mu\lambda d^{-1} \tanh(2\pi x/d); \quad (12)$$

when  $y = \frac{1}{2}d$  it is

$$\mu\lambda d^{-1} \tanh(\pi x/d). \quad (13)$$

Starting from the equilibrium configuration (a) fig. 11, for which each negative centre is midway between two positive centres on a line perpendicular to the slip planes, we shall give the positive lattice a displacement  $\xi$  relative to the negative lattice. Taking one of the positive centres as origin, the co-ordinates of the negative centres are  $(-\xi + ma, (n + \frac{1}{2})d)$  where  $n$  and  $m$  are integral numbers positive or negative,  $a$  is the distance apart of centres of dislocation on a slip plane. If  $a = d$ , the positive centres lie at the corners of a square lattice. As  $\xi$  increases from zero the shear stress at the origin due to the negative centres changes, but that due to the positive centres, which remain fixed, is unaltered. If  $S'$  is the shear stress at the origin due to all the centres of dislocation (except the positive centre which is situated there), then  $S' = 0$  when  $\xi = 0$ . The contribution to  $S'$  due to the displacement of the  $m$ th row of negative dislocations is (see (13))

$$\mu\lambda d^{-1} \{\tanh \pi d^{-1}(ma - \xi) - \tanh \pi d^{-1} ma\}. \quad (14)$$

As  $m$  increases the value of this expression rapidly decreases, so that a convergent expression is obtained for  $S'$ , namely,

$$S' = \mu\lambda d^{-1} \sum_{m=-\infty}^{+\infty} \{\tanh \pi d^{-1}(ma - \xi) - \tanh \pi d^{-1} ma\}. \quad (15)$$

If the change in  $S'$  is traced as  $\xi$  increases, it will be found that each increase in  $\xi$  by an amount  $a$  gives rise to a change in  $S'$  of amount  $-2\mu\lambda d^{-1}$ . This change may be regarded as being due to the separation of a layer of positive centres at the right-hand boundary of the crystal and a layer of negative centres on the left, thus it is a polarization effect arising from the distribution of dislocations on the boundary of the crystal. It will be noticed that the effect of this polarization is to increase the shear stress at all points in the field, not only at the lattice points, but an amount  $-2\mu\lambda/d$  for each increase of amount  $a$  in the displacement. Thus the mean shear stress over the whole crystal due to the presence of the centres of dislocation is increased by  $-2\mu\lambda/d$ .

Besides the effect of the relative displacement of the two lattices of centres of dislocation the conditions of stress at the boundary of the crystal itself have an unknown effect on the stress within it. If we suppose that the mean stress in any portion of the crystal is equal to  $S$  the externally applied stress, *i.e.*, if we suppose that the mean stress taken over a portion of the crystal which includes a large number of centres of dislocation is uniform throughout the volume of the crystal and equal to  $S$ , then the mean stress due to the combined effects of the distribution of centres of dislocation and to the conditions at the boundary must be zero. Thus to obtain the true shear stress at positive centres of dislocation, due to the displacement of the lattice of negative centres, we must subtract from the expression (15) the mean value of the shear stress in the crystal due to the whole body of the centres of dislocation.

### *Mean Value of Shear Stress.*

The mean value of the shear stress due to the distribution of centres of dislocation may be found by taking the mean value over a strip parallel to the axis of  $y$  and of breadth  $a$ . If this mean stress be represented by  $S_M$ , the change in  $S_M$  which occurs when the displacement of the negative centres relative to the positive ones (which we suppose fixed) increases by  $\delta\xi$  may be regarded as being due to the removal of a strip of thickness  $\delta\xi$  on one side of the strip and replacing it by a strip also of the thickness  $\delta\xi$  but located in a region where the shear stress has increased by  $-2\mu\lambda/d$ . The mean value of  $S_M$  is therefore increased by  $-2\mu\lambda\delta\xi/ad$ . Hence

$$\left. \begin{aligned} \frac{dS_M}{d\xi} &= -\frac{2\mu\lambda}{ad}, \\ S_M &= -2\mu\lambda\xi/ad \end{aligned} \right\}. \quad (16)$$

so that

Hence when the effects of the boundary have been taken into consideration in such a way that the mean stress throughout the body of the crystal is equal to the externally applied stress the resultant stress at a positive centre is  $S + (S' - S_M)$ .

If the temperature is greater than  $T_D$  so that centres of dislocation move freely under the influence of the slightest shear stress, the equilibrium condition is therefore

$$S + S' - S_M = 0, \quad (17)$$

or combining (15), (16), and (17)

$$S = \mu \lambda d^{-1} (F), \quad (18)$$

where

$$F = -\frac{2\xi}{a} + \sum_{m=-\infty}^{+\infty} \left\{ \tanh \frac{\pi m a}{d} - \tanh \frac{\pi}{d} (m a - \xi) \right\}. \quad (19)$$

$F$  is a function of  $\xi/a$  and of  $a/d$ . The values shown graphically in the curves of fig. 12 have been calculated for values of  $\xi/a$  ranging from 0 to 0.5 for each of the three lattices  $a = d$ ,  $a = 0.8d$ ,  $a = 0.5d$ . It will be seen that each

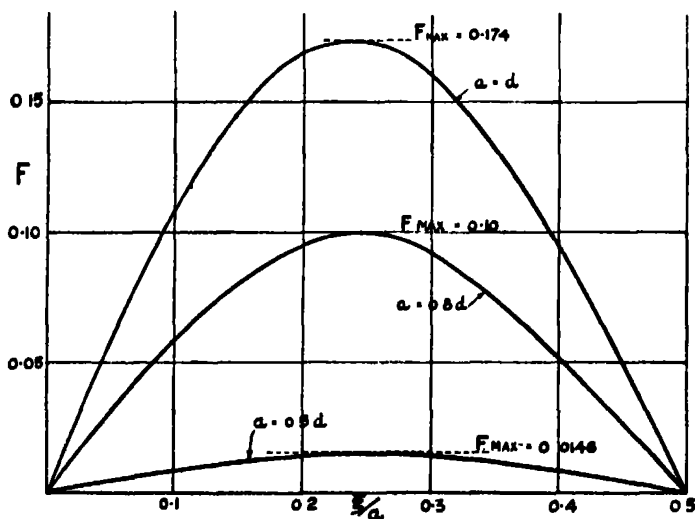


FIG. 12.

curve rises from zero at  $\xi = 0$  and falls to zero at  $\xi/a = 0.5$ . This is to be expected because  $\xi/a = 0.5$  corresponds to the equilibrium arrangement represented by (b), fig. 11. The fact that  $F$  is positive over the whole range from  $\xi/a = 0$  to 0.5 shows that (b) is an unstable configuration while (a) is stable. In each case  $F$  has a maximum value (see fig. 12). For  $a = d$  it is  $F_{\max} = 0.174$ . For  $a = 0.8d$  it is  $F_{\max} = 0.10$ . For  $a = 0.5d$  it is  $F_{\max} = 0.0146$ .



It will be seen, therefore, that when the externally applied shear stress  $S$  is less than  $\mu\lambda d^{-1} (F_{\max})$  a value of  $\xi$  can be found such that every centre of dislocation, positive or negative, is in a position where the shear stress is zero, so that it can remain at rest. When  $S$  rises above this value the positive centres will migrate steadily to the right and the negative ones to the left, and plastic distortion will set in. Thus the crystal has a finite shear strength

$$S = \mu\lambda d^{-1} (F_{\max}). \quad (20)$$

If the crystal is initially perfect, the first few dislocations migrate through it under the action of a shear stress which may be regarded as infinitesimally small. As the distortion proceeds, however, the number of dislocations will increase and the average value of  $d$  will decrease so that the resistance to shear will increase with the amount of distortion. This explanation of hardening by cold work is identical in principle with one put forward in a qualitative form by the present writer some years ago,\* but the introduction of the idea of a unit dislocation which gives rise to a calculable field of elastic stresses has now made it possible to present the theory in a more definite and constructive form than formerly.

#### Diagonal Lattices of Dislocations.

So far the only arrangement of lattices which has been discussed is that in which the two kinds of dislocation lie on two interpenetrating rectangular nets. Another equilibrium arrangement is that in which the positive and negative dislocations lie on interpenetrating lattices such that each cell is a

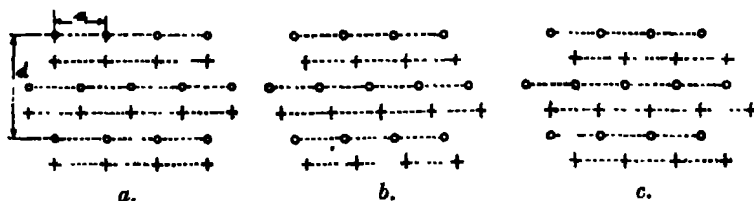


FIG. 13.—Equilibrium arrangements of dislocations, (a) and (c) stable, (b) unstable.

rhombus with one of its diagonals parallel to the slip planes. A diagonal arrangement of this kind is shown in fig. 13. If the dislocations are initially arranged as configuration (a), fig. 13, a relative displacement of the positive and negative lattices through  $\frac{1}{2}a$  brings the system to configuration (c) which is the mirror image of (a) in the slip planes. Thus if (a) is stable (c) will also be stable and *vice versa*.

\* 'Trans. Faraday Soc.,' *loc. cit.*

The system must also be in equilibrium when the displacement is  $0.25a$  and  $0.75a$ . This configuration is shown in (b), fig. 13.

If  $a$  and  $d$  represent the two diagonals of the lattice cell and  $\xi$  the displacement of the positive relative to the negative lattice from the equilibrium positions (a), fig. 13, the shear stress at one of the positive centres due to a row of negative centres is (see equation (12))

$$\frac{\mu\lambda}{d} \tanh \frac{2\pi}{\lambda} \left( \frac{ma}{2} - \xi \right),$$

where  $m$  takes all integral values. Proceeding on the same lines as in the case of the rectangular lattice the condition of equilibrium is

$$S' = \mu\lambda d^{-1} (F), \quad (21)$$

where

$$F = -\frac{4\xi}{a} + \sum_{m=-\infty}^{m=+\infty} \left\{ \tanh \frac{2\pi}{d} \left( \frac{na}{2} \right) - \tanh \frac{2\pi}{d} \left( \frac{na}{2} - \xi \right) \right\}. \quad (22)$$

When  $a = d$ , i.e., when a square lattice is set diagonally relative to the slip planes, the following values for  $F$  were calculated from (22)

$\xi/a$ .....	0	0.05	0.1	0.15	0.20	0.25
$F$ .....	0	0.109	0.169	0.160	0.095	0

The maximum value is  $F_{\max} = 0.17$ .

The fact that  $F$  is positive in the range  $\xi/a = 0$  to  $\xi/a = 0.5$  shows that the equilibrium configurations (a) and (c), fig. 13, are stable while (b) (for which  $\xi/a = 0.25$ ) is unstable.

When  $d = a\sqrt{3}$  the centres of dislocation lie on a net consisting of equilateral triangles. The corresponding values of  $F$  calculated from (22) are:—

$\xi/a$ .....	0	0.05	0.10	0.15	0.20	0.25
$F$ .....	0	0.0179	0.0289	0.0283	0.0176	0

and  $F_{\max} = 0$ .

#### *Relationship between Strain and the Motion of Centres of Dislocation.*

In the simple case when only one centre of dislocation is present the effect of moving it across the crystal from left to right has already been seen, fig. 4. The part of the crystal above its path is displaced through a distance  $\lambda$  relative to the lower part. If the centre of dislocation does not travel across the crystal from side to side but merely moves through a distance  $x$  from the outer surface, the displacement is  $x\lambda/L$ , where  $L$  is the total length along the slip plane

through which a dislocation can move.  $L$  might, as we have so far assumed, be the total width of the specimen or it might be the width of any subdivision of the crystal the boundary of which is impervious to centres of dislocation. The total shift due to centres regularly spaced at intervals  $a$  along the slip plane is therefore  $L\lambda/2a$ . The shift due to a row of positive centres, which have moved into the crystal from the left, is the same as that due to a row of negative centres which have moved in from the right.

Since in the rectangular arrangement, fig. 11, the operative slip planes are spaced at distance  $\frac{1}{2}d$  apart the relative shift of two slip planes separated by a thickness  $h$  is therefore  $hL\lambda/ad$ . By definition therefore the shear strain

$$s = L\lambda/ad. \quad (23)$$

For the diagonal arrangement of dislocations, fig. 13, the operative slip planes are situated at distance  $\frac{1}{2}d$  apart so that

$$s = 2L\lambda/ad. \quad (24)$$

### *Theoretical Stress-strain Relationship.*

We have seen that the least shear stress at which plastic flow occurs when the temperature is higher than  $T_D$  is  $S = \mu\lambda d^{-1}F_{\max}$ . Combining this with (23) or (24)  $d$  can be eliminated. The resulting equation is

$$\frac{S}{\mu\sqrt{s}} = \kappa \sqrt{\frac{\lambda}{L}}, \quad (25)$$

where  $\kappa$  is a constant depending on the particular arrangement of centres of dislocation. The  $(S, s)$  curve is therefore parabolic. The values of  $\kappa$  are given below.

Rectangular lattice .....	$a = d$	$\kappa = 0.174$
	$a = 0.8d$	$\kappa = 0.089$
	$a = 0.5d$	$\kappa = 0.010$
Diagonal lattice .....	$a = d$	$\kappa = 0.120$
	$a = d/\sqrt{3}$	$\kappa = 0.016.$

If the temperature is below  $T_D$ , all the foregoing work is unchanged, but the condition of equilibrium of a centre of dislocation is now

$$S - S_T = \mu\lambda d^{-1}(F),$$

where  $S_T$  is the finite stress which must be attained before a single dislocation

can migrate in a perfect crystal. The resulting strain hardening or plastic stress-strain relationship is then

$$\frac{S - S_T}{\mu \sqrt{s}} = \kappa \sqrt{\frac{\lambda}{L}}. \quad (26)$$

*Choice of a Value for  $\kappa$  for Comparison with Experiment.*

It will be seen that comparatively small variations in the arrangements of centres of dislocation produce very large changes in the value of  $\kappa$ . A crowding of positive centres of dislocation on one set of slip planes increases the ease with which negative centres can migrate along the intermediate planes. Thus in the rectangular lattice of dislocations a decrease in the ratio  $a/d$  corresponds with a much greater decrease in  $\kappa$ .

In trying to form some estimate as to which of these regular arrangements of dislocations is likely to represent most closely the true condition of a plastically strained crystal it is necessary to consider the manner in which the dislocations may arise at the boundary. Since in this theory the crystal is undisturbed by the passage of a dislocation, i.e., the crystal is just as perfect behind it as it is in front, the dislocations might be supposed to arise at the boundary or within the crystal owing to thermal agitation in a random manner. The hypothesis that the points of origin of the dislocations are distributed at random precludes the formation of regular lattices of dislocations of the type we have been considering. The kind of regularity which produces very low values of  $\kappa$  is therefore unlikely to arise.

On the other hand, it is difficult to calculate the effect of a random production of dislocations because the mutual action of neighbouring centres of dislocation must itself tend to prevent the formation of a purely random distribution of centres inside the crystal. If, for instance, in the course of plastic distortion a positive and a negative centre approach one another from opposite sides of the crystal along planes which are separated by a very small distance  $h$ , they will not separate again till  $S$  reaches the value  $\mu\lambda/2\pi h$ . Thus a pseudo regularity in which the distance of a dislocation from its nearest neighbour tends always to be about  $\mu\lambda/2\pi S$  is likely to arise.

We might, perhaps, visualize the whole process of plastic deformation as follows. At points distributed at random in a crystal slipping begins, owing possibly, to thermal agitation, by separation of a positive and negative dislocation. These move away from one another to an average distance apart  $L$  under the action of the external shear stress  $S$ . Slipping through one lattice

distance has therefore occurred along each line of length  $L$  which joins a positive and negative pair of centres. If at any stage there are  $N$  such pairs of dislocation lines per unit volume, the average distance of the nearest negative dislocation from any given positive one is proportional to  $1/\sqrt{N}$ . A very rough approximation is  $1/2\sqrt{N}$ .

If only the shear stress due to the nearest negative dislocation is considered, the shear stress necessary for its escape is of order of magnitude

$$S = \frac{\mu\lambda}{2\pi} (2\sqrt{N}) = \mu\lambda\sqrt{N}/\pi.$$

The strain is evidently  $s = N\lambda L$  so that

$$\frac{S}{\mu\sqrt{s}} = \frac{1}{\pi} \sqrt{\frac{\lambda}{L}}. \quad (27)$$

This expression is identical with (25) if  $\kappa = \pi^{-1} = 0.32$ , but it is likely to overestimate  $S$  because it takes account only of the effect of the nearest dislocation of opposite sign.

In choosing a value to assume for  $\kappa$  in order to compare the theory with observation we might take it that there are reasons for supposing that  $\kappa = 0.1$  is too low, and  $\kappa = 0.3$  is too high, accordingly we may use  $\kappa = 0.2$  as the most probable value in the present very imperfect state of the theory. Calculations based on this value should at any rate give the correct order of magnitude for the strength of crystals.

---

*The Mechanism of Plastic Deformation of Crystals.*  
*Part II.—Comparison with Observations.*

By G. I. TAYLOR, F.R.S., Royal Society Yarrow Professor.

(Received February 7th, 1934.)

*Comparison with Observed Plastic Stress-strain Relationships.*

According to the theory given in Part I the strain-hardening or plastic stress-strain curve for a pure metal should be a parabola. In figs. 1, 2, and 3, Part I, parabolas are drawn, the parameters being chosen so that they lie as close as possible to the points which represent actual observations. It will be seen that for aluminium and gold the agreement is good. For a single crystal of copper the agreement is not good, but, on the other hand, the plastic stress-strain curve for polycrystalline specimens of copper which is shown in fig. 1 is very nearly parabolic over a large range.

The observations for iron seem to show that there is a small finite elastic limit, i.e.,  $S_T$  may be finite. Parabolas corresponding with the existence of a small elastic limit and with no elastic limit have been drawn. It seems that the observed points lie rather closer to the former curve. In any case, the observed curves have the essential characteristic of the theoretical ones that they are very steep at small strains, but get less and less steep as the strain increases.

In some branches of physics the measure of agreement between theory and observation which is shown in figs. 1, 2, and 3, Part I, would perhaps not be considered encouraging. It must be remembered, however, that up to the present no theory of the strength of metals has been devised which is capable of being expressed in an analysable form. Also the plastic stress-strain curves of similar metals vary so much in detail that it would be impossible to devise any general theory which would account so accurately for the behaviour of every metal crystal that good agreement with actual individual observations could be attained. Another difficulty is that the large changes which very small amounts of impurity can cause in the stress-strain curves make it difficult to be certain that the observed curves are really those which would be found if the test were made with a perfect crystal of a metal containing no impurities.

*Necessity for a Mosaic Structure of System of Faulting.*

The stress-strain relationship (25) Part I contains the ratio  $\lambda/L$ . It is clear that  $L$  must not be interpreted as an external linear dimension of a single crystal specimen because the observed stress-strain relationship does not depend on the size of the specimen employed. The length  $L$  must be a length connected with the structure of the crystal. It represents the distance through which a

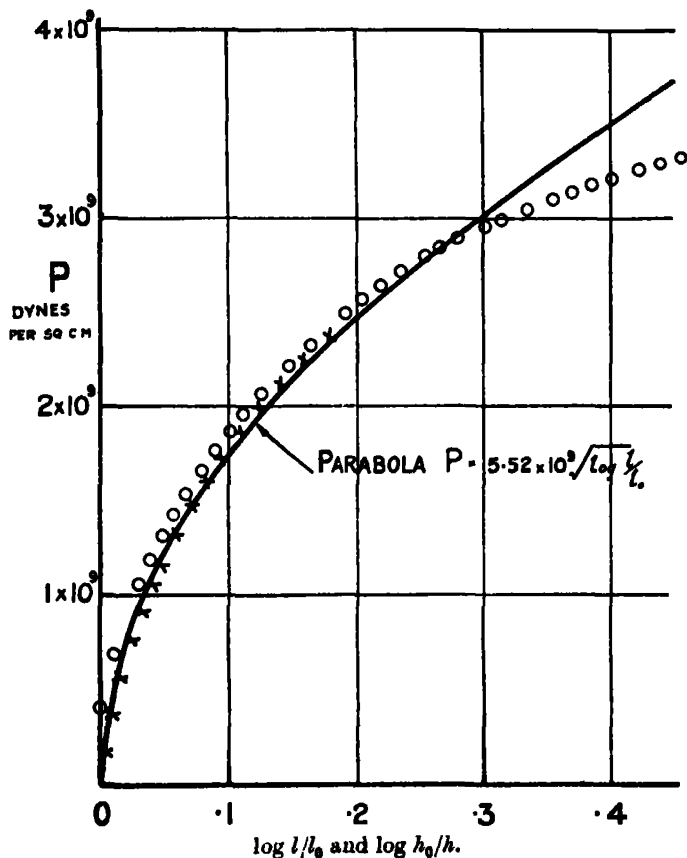


FIG. 1.—Plastic stress-strain curve for polycrystalline copper in direct tension or compression.  $\times$  tension,  $\log l/l_0$ ;  $\circ$  compression,  $\log h_0/h$ .

dislocation can travel freely along a slip plane under the influence of a small shear stress before being held up by some fault in the perfection of the regular crystal structure, or surface of misfit. It is therefore a linear dimension connected with the spacing of such faults in the crystal. If the faults or surfaces of misfit are everywhere opaque to dislocations,  $L$  would be the linear dimension of a cell of a superstructure or a mosaic.

From the parabolas which most nearly fit the observed plastic stress-strain curves the values of  $S/\sqrt{s}$  may be found. Thus for aluminium the parabola shown in fig. 1, Part I, is  $S = 3.8 \times 10^8 \sqrt{s}$  dynes per square centimetre.  $\mu$  may be taken as  $2.6 \times 10^{11}$  so that  $S/\mu \sqrt{s} = 1.46 \times 10^{-3}$ . For copper the parabola of fig. 2, Part I, is  $S = 8.8 \times 10^8 \sqrt{s}$  and  $\mu = 4.55 \times 10^{11}$ , so that  $S/\mu \sqrt{s} = 1.94 \times 10^{-3}$ . For gold the parabola of fig. 2, Part I, is  $S = 4.5 \times 10^8 \sqrt{s}$  and  $\mu = 2.8 \times 10^{11}$ , so that  $S/\mu \sqrt{s} = 1.6 \times 10^{-3}$ . For iron  $\mu = 8.3 \times 10^{11}$  accordingly the parabola  $S = 2 \times 10^9 \sqrt{s}$  of fig. 3, Part I, which corresponds with the assumption that iron has no finite elastic limit gives  $S/\mu \sqrt{s} = 2.4 \times 10^{-3}$ . The parabola  $S = 1.2 \times 10^8 + 1.5 \times 10^9 \sqrt{s}$ , which corresponds with the assumption that iron crystals have an elastic limit of  $1.2 \times 10^8$  dynes per square centimetre, gives  $(S - S_T)/\mu \sqrt{s} = 1.8 \times 10^{-3}$ .

Table I.

(1)	(2)	(3)	(4)	(5)	(6)
	$S/\mu \sqrt{s}$ .	$b$ (cm.).	$\lambda$ (cm.).	$L\kappa^{-2}$ (cm.).	$L$ (cm.). ( $\kappa = 0.2$ ).
Al .....	$1.46 \times 10^{-3}$	$4.05 \times 10^{-3}$	$2.86 \times 10^{-3}$	$1.34 \times 10^{-2}$	$5.3 \times 10^{-4}$
Cu .. ..	$1.94 \times 10^{-3}$	$3.6 \times 10^{-3}$	$2.55 \times 10^{-3}$	$0.68 \times 10^{-2}$	$2.7 \times 10^{-4}$
Au .. .	$1.61 \times 10^{-3}$	$4.08 \times 10^{-3}$	$2.87 \times 10^{-3}$	$1.11 \times 10^{-2}$	$4.4 \times 10^{-4}$
Fe .. .	$2.4 \times 10^{-3}$	$2.86 \times 10^{-3}$	$2.47 \times 10^{-3}$	$0.43 \times 10^{-2}$	$1.7 \times 10^{-4}$

(Observed spacing of surface marks in crystals :—

Bismuth .....	$1.4 \times 10^{-4}$ cm.	(Goetz).
Zinc .. .	$0.8 \times 10^{-4}$ cm.	(Straumanis).
Iron .. .	$0.2b \times 10^{-4}$ cm.	(Belaiew).

These values which are set forth in column 2 of Table I are according to the present theory equal to  $\kappa\sqrt{\lambda/L}$ . In this expression  $\lambda$  is the least distance through which the atoms above the slip plane must be displaced relative to those below in order that a perfect crystal structure may be re-formed. With face-centred cubic crystals like copper, aluminium, and gold, for which the possible directions of slip are parallel to the diagonals of faces of the cube,  $\lambda = b/\sqrt{2}$ , where  $b$  is the length of the side of the unit cube. For iron which is body-centred, the direction of slip is parallel to a cube diameter and  $\lambda = \frac{1}{2}b\sqrt{3}$ . The values of  $b$  are known from X-ray data and are given in column 3 of Table I, the corresponding values of  $\lambda$  being given in column 4. Dividing the figures in column 4 by the squares of the figures given in column 2 values are found for  $L\kappa^{-2}$ ; these are given in column 5.



It is only possible to proceed further by assuming a value for  $\kappa$ . Taking  $\epsilon = 0.2$  the values given in Table I, column 6, are found for  $L$ .

*Comparison with Observations of Systems of Faulting and Lineage Systems in Metals.*

The length  $L$  appears in this theory as the distance which a dislocation can travel before being stopped by some barrier such as a surface of misfit. The existence of such surfaces inside a crystal is essential to the theory, but so far they have been regarded as hypothetical. On the other hand, a large number of observations appear to show that surfaces of misfit do occur in crystals.

By suitable treatment Goetz has shown that it is possible to reveal on the surface of bismuth crystals some very closely spaced marks. These consist of three sets of lines spaced  $1.4 \times 10^{-4}$  cm. apart forming a triangular pattern.

When zinc or cadmium\* are deposited from vapour they form a system of plates  $0.8 \times 10^{-4}$  cm. thick.

Etching pits in iron† seem to point to the existence of a block structure with sides of length  $0.25 \times 10^{-4}$  cm.

A micro photograph‡ which is referred to by Zwicky shows a structure which is brought out by etching a copper crystal. This structure indicates a rectangular arrangement of surfaces of misfit (or according to Zwicky a superstructure) the fundamental linear dimension of which appears (judging from the photograph) to be about  $1.5 \times 10^{-4}$  cm. These figures are given at the end of Table I in order to facilitate comparison with the values of  $L$  which are necessary in the foregoing theory to account for the observed plastic stress-strain relationships.

It will be seen that the observed spacings are of the same order of magnitude as the values of  $L$  given in column 6, Table I, namely,  $10^{-4}$  cm., but  $L$  is always rather greater than any of the smallest spacings which have been observed. This discrepancy might perhaps be due to an error in calculating the value of  $\kappa$ . If  $\kappa$  were 0.1 instead of 0.2, all the figures in column 6 would be a quarter of the values there given, and the values of  $L$  would be very close to the observed spacings, i.e., about  $1.0 \times 10^{-4}$  cm.

This explanation of the discrepancy would involve the assumption that the surfaces of misfit which are observed are completely opaque to the passage of

\* Straumanis, 'Z. Physik,' vol. 13, p. 316 (1931); vol. 19, p. 63 (1932).

† Belaiew, 'Proc. Roy. Soc.,' A, vol. 108, p. 295 (1925).

‡ Photograph, Plate 1, "Brown Boveri Rev.," January, 1929.

dislocations over their whole area. It seems unlikely that such an hypothesis can be true. Whatever conception is held as to the nature of surfaces of misfit it seems that the disturbance of the lattice must be greater in some parts of them than in others. Consider, for instance, Darwin's type of mosaic in which the whole crystallite consists of blocks or bands in which the orientation of the crystal axes of any one block differs from that of its neighbours by several minutes of arc. Fig. 2 represents two neighbouring blocks, bands, or dendrites

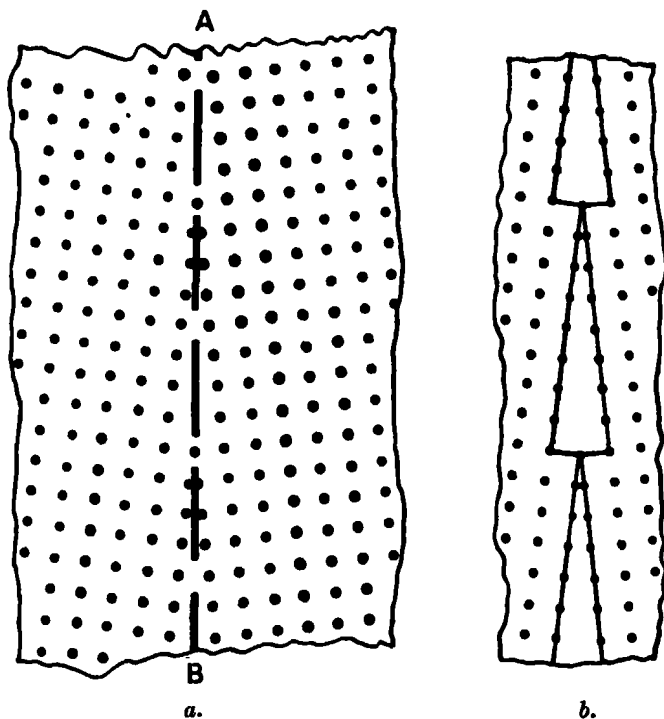


FIG. 2.—Boundary of two crystals at slightly different orientations : (a) showing portions of surface of misfit opaque to dislocations ; (b) showing boundaries of separate crystal portions.

of a crystal with a cubic lattice. The surface of misfit is shown in fig. 2 (a) as AB. If the positions of the atoms on the two sides of AB be examined, it will be seen that in certain regions, some of which are represented in fig. 2 (a) by gaps in the line AB, the distance of atoms on one side from the nearest atoms on the other is the same as that which belongs to the perfect crystal structure. At these points therefore the disturbance of the lattice in passing from one block to the other is small. At the intermediate points the disturbance is a maximum. It is to be expected that as the temperature increases

the proportion of the whole area of the surfaces of misfit which are transparent to dislocations will increase.

If  $Z$  represents the proportion which is opaque to dislocations, it is to be expected that  $Z$  will decrease with rising temperature. Thus at very low temperatures  $Z$  might be 1.0 while at a sufficiently high temperature it might tend to 0. When  $Z = 1.0$  and the surfaces of misfit are entirely opaque to dislocations the length  $L$  is equal to  $B$ , the distance between surfaces of misfit. Thus we might expect  $L$  to approximate to  $B$  at low temperatures, but at higher temperatures  $L$  should be greater than  $B$ .

### *Variation in Plasticity with Temperature.*

It has been seen already that at room temperatures the calculated values of  $L$  are of the same order as, but greater than, the observed spacings  $B$  of surfaces of misfit. It remains to be seen whether as the temperature is reduced  $L$  tends to become equal to  $B$ , as the present theory predicts it should.

A very complete set of experiments on the plasticity of a metal crystal over a wide range of temperatures has been made by Boas and Schmid,\* who determined the  $(S, s)$  curves of aluminium crystals over the range for  $-185^{\circ}\text{C}$ . to  $600^{\circ}\text{C}$ . Their results are shown in fig. 3, which is reproduced from their paper. The shaded areas cover all the observations made with many specimens at each temperature.

It will be seen that the  $(S, s)$  curves are very similar at all temperatures and agree at  $18^{\circ}\text{C}$ . with the curve of fig. 1, Part I, but that as the temperature rises they become horizontal at rapidly decreasing values of  $s$ . In order to find the value of  $S/\sqrt{s}$  from which  $L$  must be calculated the values of  $S$  have been taken from Boas and Schmid's curves at  $s = 0.2$ . These are given in column 2, Table II.

The values of  $L$  calculated from the formula  $\frac{S}{\mu\sqrt{s}} = \kappa\sqrt{\frac{\lambda}{L}}$  where  $\kappa = 0.2$  are given in column 5, Table II. It will be seen that  $L$  decreases with decreasing temperature till at  $-185^{\circ}\text{C}$ . it is  $1.8 \times 10^{-4}\text{ cm}$ . This is very close indeed to the spacing of surfaces of misfit which have been observed (see data at base of Table I).

### *Interpretation of Boas and Schmid's Results.*

According to the foregoing theory the decrease in  $S/\sqrt{s}$  with increasing temperature is due to an increase in  $L$ . This increase in  $L$  is due to a decrease

\* 'Z. Physik,' vol. 71, p. 713 (1931).

in the fraction  $Z$  of the surface of misfit which can act as barriers to the dislocations. If the surfaces of misfit are uniformly spaced at distance  $B$  apart, it is possible to calculate the mean value of  $L$  as a function of  $Z$ . Suppose that

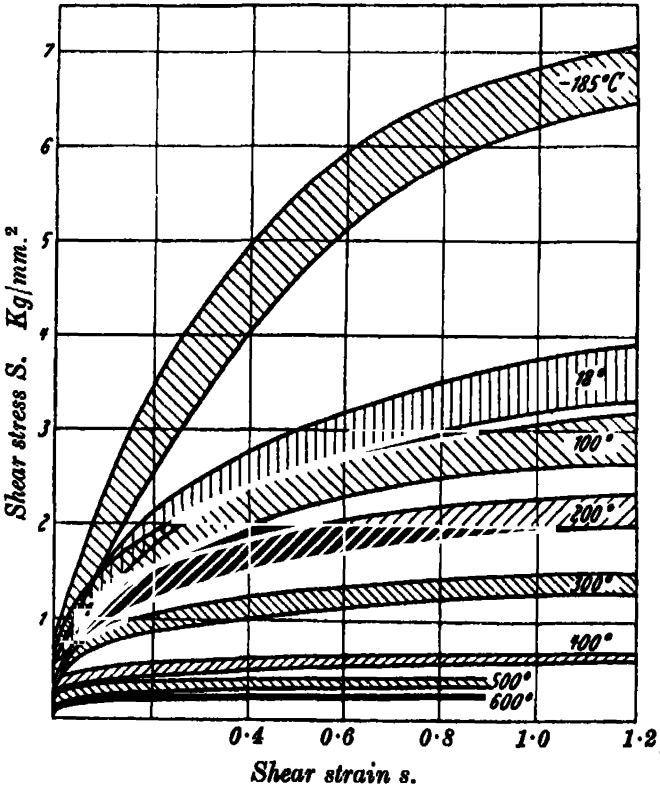


FIG. 3.—( $S, s$ ) curves for aluminium crystals (Boas and Schmid).

Table II.

Tempera- ture.	$S$ kg. mm. <sup>2</sup> .	$S/\mu\sqrt{s}$ $\times 10^3$ .	$L\kappa^{-2}$ $\times 10^3$ .	$L$ (cm.) ( $\kappa = 0.2$ ).	$Z = \frac{2B}{L+B}$ .		
					$B = 1.5$ $\times 10^{-4}$ cm.	$B = 1.0$ $\times 10^{-4}$ cm.	$B = 0.5$ $\times 10^{-4}$ cm.
° C.							
-185	3.0	2.53	0.477	$1.8 \times 10^{-4}$	0.91	0.71	0.54
18	1.95	1.64	1.06	$4.2 \times 10^{-4}$	0.53	0.38	0.21
100	1.72	1.45	1.36	$5.4 \times 10^{-4}$	0.43	0.31	0.17
200	1.37	1.156	2.14	$8.6 \times 10^{-4}$	0.30	0.21	0.11
300	0.97	0.818	4.28	$1.7 \times 10^{-3}$	0.16	0.11	0.057
400	0.50	0.472	12.84	$5.1 \times 10^{-3}$	0.06	0.038	0.019
500	0.34	0.287	34.8	$1.4 \times 10^{-2}$	0.02	0.014	0.007
600	0.20	0.169	100.0	$4.0 \times 10^{-2}$	0.01	0.005	0.002

each surface of misfit acts as an independent barrier to a fraction  $Z$  of the centres of dislocation which approach it. We may consider a large number  $N$  of path of centres of dislocation which cross a plane  $AD$ , fig. 4, which is half-way between two surfaces of misfit and we may calculate their total length. Taking

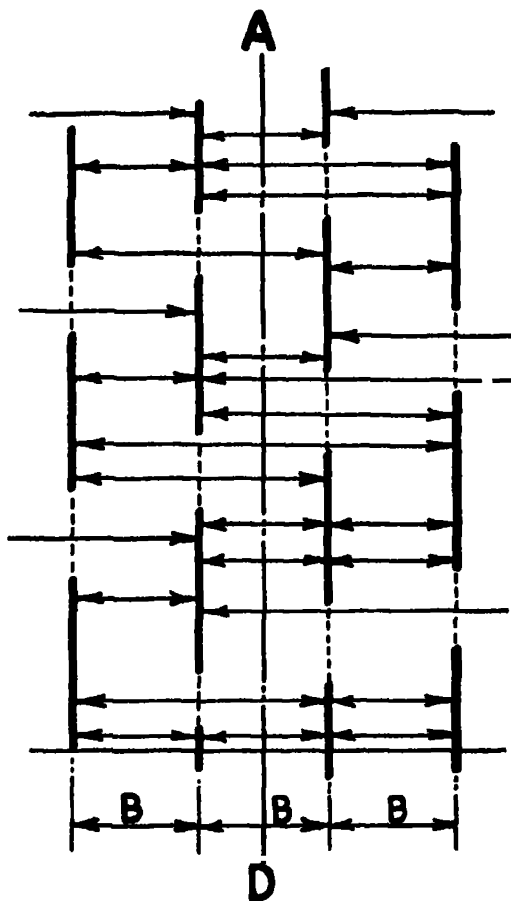


FIG. 4.—Paths of dislocations ending on the opaque portions of surfaces of misfit.

first all the parts of the paths which lie to the left of the plane,  $NZ$  of them end on the first barrier and the total length of these is  $\frac{1}{2}BNZ$ . Of the  $N(1-Z)$  paths which pass the first barrier  $NZ(1-Z)$  end on the second barrier and their total length is  $NZ(1-Z)\frac{3}{2}B$ . The total length of all the paths to the left of the central plane is therefore

$$NBZ \left\{ \frac{1}{2} + \frac{3}{2}(1-Z) + \frac{5}{2}(1-Z)^2 + \dots \right\},$$

and since the total length of the parts of the paths to the right of the central plane is the same as that to the left, the total length of all  $N$  paths is

$$NBZ \sum_0^{\infty} (2n+1) (1-Z)^n,$$

which on summation is found to be  $NB \left( \frac{2-Z}{Z} \right)$ . The average length  $L$  is found by dividing this by  $N$ .

Thus

$$L = B \left( \frac{2-Z}{Z} \right), \quad (28)$$

or

$$Z = \frac{2}{(1 + L/B)}. \quad (29)$$

The expression (29) may be used to calculate  $Z$  if  $L$  and  $B$  are known. No observations for aluminium indicating the spacing or even the existence of a mosaic system appear to have been made. On the other hand, the observations of surface markings which have been made on other crystals mostly indicate that spacings are usually between  $0.5$  and  $2.0 \times 10^{-4}$  cm. It is instructive to calculate the values of  $Z$  which would correspond to surfaces of misfit separated by distances  $B = 1.5 \times 10^{-4}$ ,  $1.0 \times 10^{-4}$ , and  $0.5 \times 10^{-4}$  cm. These are given in columns 6, 7 and 8 of Table II, and are shown in the curves of fig. 5.

Inspection of these curves reveals a very striking fact. Whatever value is chosen for  $B$  the proportion  $Z$  of the surfaces of misfit which is opaque to dislocations decreases in a regular manner till it reaches a very small value at  $400^\circ$  C. If no measurements had been made above  $400^\circ$  C., the curves would naturally be extrapolated through the small extension, shown dotted in fig. 5, which would have made them cut the axis  $Z = 0$  at about  $480^\circ$  C. Thus the prediction might have been made that at temperatures above  $480^\circ$  C. the crystal would have no permanent strength so that flow would occur even at the lowest stresses. The fact that recrystallization begins about this temperature indicates that thermal agitation is capable of causing switches among the atoms from one stable position to another even at the most unfavourable points, i.e., where the misfit between neighbouring crystals is most pronounced. It is to be expected therefore that at temperatures above  $480^\circ$  C. surfaces of misfit will not be able to act as barriers to dislocations in any part of their areas. This is equivalent to the statement that at the recrystallization temperature theory would lead one to expect that  $Z = 0$ .

The fact that Boas and Schmid observed finite strengths at 500° and 600° C. might be explained either by the existence of a few foreign atoms as impurities which would still act as barriers to dislocations after the surfaces of misfit had ceased to do so, or alternatively it may be supposed that the strengths observed are not permanent strengths, but that the material is really flowing all the time and that if the velocity of deformation had been very much smaller, very much smaller strengths would have been observed. The latter view is confirmed by the fact that over the whole range of temperature from -185° C. to 400° C.

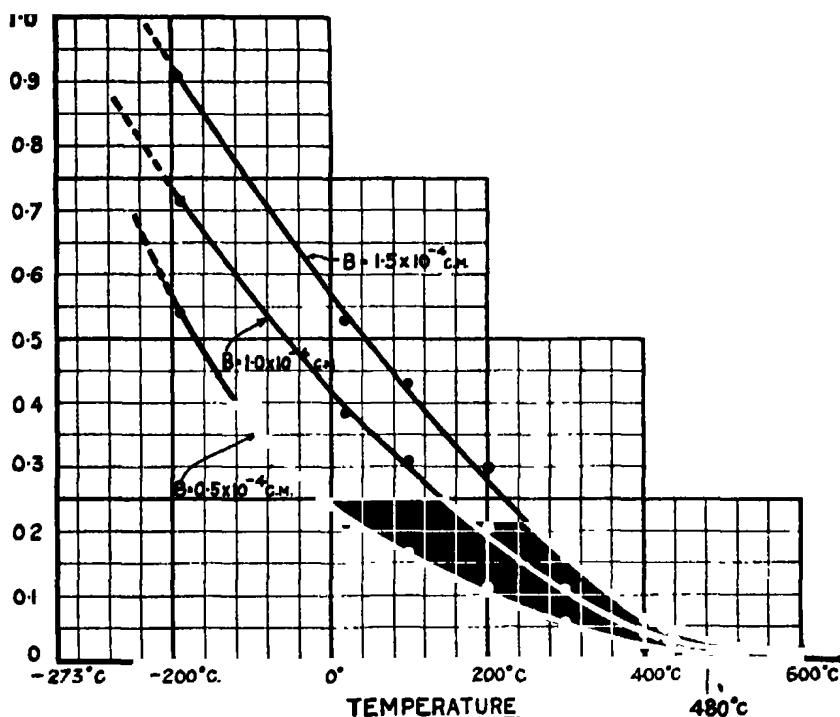


FIG. 5.—Proportion  $Z$  of surfaces of misfit which are opaque to dislocations as functions of temperature.

the distortion was found by Boas and Schmid to be of the same nature as that originally found at room temperatures\*; a slipping parallel to one plane of type (111) in direction of type [110]. At 500° C. and 600° C. the distortion was found to be of quite a different type. Test pieces pulled at these temperatures appeared to slip on several planes simultaneously.

The cross-section of a single crystal of aluminium originally circular becomes elliptical on being pulled. At temperatures below 400° C. the major axis of

\* Taylor and Elam, 'Proc. Roy. Soc.,' A, vol. 102, p. 643 (1923).

the ellipse remains almost unchanged during the elongation of the specimen, all the contraction occurring in the minor axis. At 500° C. and 600° C., on the other hand, Boas and Schmid found that both the major and minor axes contract as they would if there were no permanent resistance to shearing on any crystal plane. If, at temperatures above 500° C., slipping can occur on all crystal planes of type (111) it might be expected that the rate of slipping would be greatest in that direction and on that plane which would function alone as the slip plane at temperatures below 400° C.

The final position of the axis of the specimens in relation to the crystal axes was found by Boas and Schmid by means of X-rays. Below 400° C. it lay in the direction of type [112] which necessarily results when slipping occurs only on the plane of type (111) for which the component of shear stress is greatest. Above 500° C. they found that the end position tended to the directions [111] or [100]. Either of these directions might result from slipping on all four planes of type (111), but in order to calculate the actual movement of the direction of the axis of the specimen in relation to the crystal axes it would be necessary to make some assumption about the relationship between the rate of slipping on any plane and the component of shear stress parallel to it.

### *Plasticity at very Low Temperatures.*

According to the present theory the increase in strain hardening with decreasing temperature is due to a decrease in  $L$ . If we regard  $L$  as being determined by the spacing of existing surfaces of misfit in the crystal, which do not depend on its temperature the decrease in  $L$  can only be attributed to increasing opacity of these surfaces with decreasing temperature to centres of dislocation, i.e., to an increase in  $Z$ .

As the temperature decreases towards absolute zero the strain hardening will increase, but the increase cannot proceed beyond the point at which  $Z = 1$ , i.e., when  $L$  is equal to  $B$ . In aluminium, for instance, the strain hardening relationship corresponding with Boas and Schmid's measurements at -185° C. is  $S/\mu\sqrt{s} = 2.53 \times 10^{-3}$ , see Table II, and this corresponds with  $L = 1.8 \times 10^{-4}$  cm. The maximum possible amount of strain hardening (corresponding with  $Z = 1$ ) is :

$$\frac{S}{\mu\sqrt{s}} = \begin{cases} 2.8 \times 10^{-3} & \text{if } B = 1.5 \times 10^{-4} \text{ cm.} \\ 3.4 \times 10^{-3} & \text{if } B = 1.0 \times 10^{-4} \text{ cm.} \\ 4.8 \times 10^{-3} & \text{if } B = 0.5 \times 10^{-4} \text{ cm.} \end{cases}$$



Thus no very great increase in strain hardening is to be expected at very low temperatures. This indeed is found to hold, for experiments on cadmium crystals and on zinc crystals in liquid helium\* have shown that neither the elastic limit nor the strain hardening relationship are appreciably different at  $1.2^{\circ}$  abs. from what they are at  $4.2^{\circ}$  abs.,  $12^{\circ}$  or  $20^{\circ}$  abs. The maximum strength which the crystal can attain after considerable straining is, however, dependent on temperature.

It must be remembered that the theory accounts only for the increase in strength with increasing distortion, it is not explicitly concerned with the question whether or not the dislocations can migrate freely along slip planes without the application of a finite shear stress.

The most careful observations all seem to indicate that a finite shear stress is necessary before plastic deformation begins, but experiments at low temperatures with aluminium, zinc, and cadmium indicate that this elastic limit only changes slightly with temperature. If the observed elastic limit were really the shear stress necessary to make a single dislocation migrate in a perfect crystal, a large variation with temperature might be expected.

On the other hand, if the centres of dislocation can move freely even at the absolute zero of temperature in the absence of shear stress, the existence of a finite elastic limit might be accounted for by supposing that the crystal is not initially in a stress-free state. If surfaces of misfit exist in a crystal, they must necessarily give rise to internal stresses in the body of the crystal and these stresses would, according to the present theory, prevent the free motion of centres of dislocation provided they were greater than the externally applied stress. These considerations, however, must be regarded as speculative until some method can be found for calculating the internal stresses which might be expected.

#### *Possible Explanation of the Elastic Limit.*

If the explanation of the elastic limit just given be accepted, it may be possible to calculate at any rate its order of magnitude. According to the present theory plastic distortion begins when centres of dislocation begin to move through the crystal. If the centres will migrate under the action of any shear stress, however small, which is directed parallel to the slip planes, plastic distortion will begin when the applied shear stress is greater than the shear stresses, which occur in the body of the crystal owing to the conditions at the

\* Meissner, Polanyi, and Schmid, 'Z. Physik,' vol. 66, p. 477 (1930).

surfaces of misfit, for that is the condition that the centres can migrate through the initial "stress barriers" which must fill the body of the crystal.

*Calculation of Stresses due to Surface of Misfit.*

It is possible to form a rough estimate of the magnitude of the "stress barriers" when the surface of misfit is conceived to be the surface dividing two perfect crystal fragments of which the crystallographic axes are inclined to one another at a small angle  $\alpha$ . Consider the two cubic arrays of atoms shown in fig. 2 and suppose that the surface of separation between them bisects the angle  $\alpha$  between their crystallographic axes. Leaving out of consideration surface effects we may suppose that each fragment if separated from its neighbour would have a stepped surface, the "steps" occurring at intervals  $2b/\alpha$  along the interface. If the two fragments are brought into contact so that the edge of each step in the one meets a corresponding edge in the other, as shown in fig. 2 (b), we may think of each fragment as a continuous elastic solid and we may suppose that a pressure distribution is applied normally to each stepped surface so that the steps disappear, leaving a plane boundary. Symmetry ensures that the two portions will be in equilibrium under their mutual pressure at the boundary surface. The internal stresses are therefore those due to the normal pressure distributions which causes the stepped surface to become plane.

In the case illustrated in fig. 2 (b) the displacements along the surface of separation vary from  $-\frac{1}{2}b$  to  $+\frac{1}{2}b$ . This arrangement, however, is not that which gives rise to the least elastic disturbance. By bringing the edges of opposite steps into contact it will be seen that a line of atoms which passes through the two edges of the steps contains two more atoms than the corresponding row just below these edges. If the fragments are arranged so that the steps of one alternate with the steps of the other, there is only one more atom in the row which passes through an edge than there is in the corresponding row just beneath it. The displacement necessary to make the two fragments fit now varies from  $-\frac{1}{2}b$  to  $+\frac{1}{2}b$ , but the wave-length of the disturbance along the surface of separation is now only  $b/\alpha$  instead of  $2b/\alpha$ .

Taking axes  $ox$  in a plane boundary of an elastic solid,  $oy$  perpendicular to it and directed normally inwards the elastic system which is required is that which corresponds with a displacement  $v = \frac{1}{2}\alpha x$  over the range  $x = -b/2\alpha$  to  $x = +b/2\alpha$ , this displacement being repeated periodically with wave-length  $b/\alpha$ .

Consider the stress function

$$\chi = F(1 + my) m^{-2} e^{-my} \sin mx,$$

which satisfies the equation of equilibrium,\* namely,  $\nabla^2 \chi = 0$ . The corresponding stress components are

$$\left. \begin{aligned} X_x &= \frac{\partial^2 \chi}{\partial y^2} = -F(1 - my)e^{-my} \sin mx \\ Y_y &= -\frac{\partial^2 \chi}{\partial x^2} = -F(1 + my)e^{-my} \sin mx \\ X_y &= -\frac{\partial^2 \chi}{\partial x \partial y} = Fmye^{-my} \cos mx \end{aligned} \right\}, \quad (30)$$

At  $y = 0$ ,  $X_y = 0$  so that the  $\chi$  represents the effect of applying a normal stress  $-F \sin mx$  over the surface  $y = 0$ .

Using a method given by Love the displacements corresponding with (30) may be found. They are

$$\left. \begin{aligned} u &= -\frac{F}{2\mu m} \left( -\frac{\mu}{\lambda + \mu} + my \right) e^{-my} \cos mx \\ v &= \frac{F}{2\mu m} \left( my + \frac{\lambda + 2\mu}{\lambda + \mu} \right) e^{-my} \sin mx \end{aligned} \right\}, \quad (31)$$

where  $\lambda$  and  $\mu$  are the elastic constants.

Corresponding to the surface displacements

$$\left. \begin{aligned} u &= A_m \left( \frac{\mu}{\lambda + 2\mu} \right) \cos mx \\ v &= A_m \sin mx \end{aligned} \right\}, \quad (32)$$

the stress components are therefore

$$\left. \begin{aligned} X_x &= -2\mu mA_m \left( \frac{\lambda + \mu}{\lambda + 2\mu} \right) (1 - my) e^{-my} \sin mx \\ Y_y &= -2\mu mA_m \left( \frac{\lambda + \mu}{\lambda + 2\mu} \right) (1 + my) e^{-my} \sin mx \\ X_y &= 2\mu m^2 A_m \left( \frac{\lambda + \mu}{\lambda + 2\mu} \right) ye^{-my} \cos mx \end{aligned} \right\}. \quad (33)$$

Putting  $\xi = 2\pi x a/b$  the surface displacement is

$$\frac{1}{2}ax = f(\xi) = \frac{b\xi}{4\pi},$$

in the range  $-\pi < \xi < \pi$ , now in this range

$$\frac{1}{2}\xi = \sin \xi - \frac{1}{2} \sin 2\xi + \frac{1}{3} \sin 3\xi \dots$$

\* See Love, "Elasticity," 4th ed., p. 204.

so that at  $y = 0$

$$v = \frac{b}{2\pi} \sum (-1)^{n+1} \frac{\sin n\xi}{n}. \quad (34)$$

Comparing (32) with (34) it will be seen that surface displacement  $v$  may be expressed in the form  $\Sigma A_m \sin mx$  provided  $n\xi = mx$  and

$$A_m = (-1)^{n+1} b/2\pi n.$$

Hence from (33) the stress components are

$$\left. \begin{aligned} X_x &= 2\mu\alpha \left( \frac{\lambda + \mu}{\lambda + 2\mu} \right) \sum (-1)^n (1 - n\eta) e^{-n\eta} \sin u\xi \\ Y_y &= 2\mu\alpha \left( \frac{\lambda + \mu}{\lambda + 2\mu} \right) \sum (-1)^n (1 + n\eta) e^{-n\eta} \sin u\xi \\ X_y &= 2\mu\alpha \left( \frac{\lambda + \mu}{\lambda + 2\mu} \right) \sum (-1)^{n+1} n\eta e^{-n\eta} \cos n\xi \end{aligned} \right\}, \quad (35)$$

where  $\eta = 2\pi ay/b$ .

These expressions represent the distribution of stress which might be expected near a surface of misfit between two crystal fragments the crystallographic axes of which are inclined to one another at a small angle  $\alpha$ .

In order to visualize the nature of this stress distribution we may calculate the value of  $X_y$  at various points.

Let

$$f(\xi\eta) = \sum (-1)^{n+1} n\eta e^{-n\eta} \cos u\xi \quad (36)$$

so that

$$X_y = 2\mu\alpha \left( \frac{\lambda + \mu}{\lambda + 2\mu} \right) f(\xi\eta). \quad (37)$$

By summing the series (36) it can be shown that along the line  $\xi = 0$ ,

$$f(\xi\eta) = \frac{\eta e^{-\eta}}{(1 + e^{-\eta})^2}; \text{ when } \xi = \pi, f(\xi\eta) = \frac{\eta e^{-\eta}}{(1 - e^{-\eta})^2}, \text{ and when } \xi = \frac{1}{2}\pi,$$

$$f(\xi\eta) = \frac{2\eta e^{-2\eta}}{(1 + e^{-2\eta})^2}. \text{ These functions are shown graphically in fig. 6.}$$

It will be seen that when  $\eta > 1$  the order of magnitude of the shear stresses in the interior of the crystal due to surfaces of misfit is

$$X_y = 2\mu\alpha \left( \frac{\lambda + \mu}{\lambda + 2\mu} \right) \eta e^{-\eta}.$$

To find numerical values for this expression it is necessary to assume values for  $\alpha$  and  $\eta$ . If the surfaces of misfit are spaced at distances  $L$  apart, the value of  $\eta$  at the plan mid-way between two such surfaces is  $\pi\alpha L/b$ .

If  $L$  is taken as  $10^{-4}$  cm. and  $b$  as  $3 \times 10^{-8}$  cm.,  $\eta = 10^4\alpha$ . Taking  $\alpha = 1' = 3 \times 10^{-4}$  radians, i.e., assuming that the crystal consists of a Darwin

type of mosaic in which the orientation of the crystal axes of neighbouring fragments differs on the average by  $1'$  of arc,  $\eta = 3$  and  $\eta e^{-\eta} = 0.15$ . Thus the order of magnitude of  $X_v$  is  $2\mu(3 \times 10^{-4})(0.015) = 10^{-4} \mu$ . For aluminium  $\mu = 2.6 \times 10^{11}$ , so that

$$X_v = 2.6 \times 10^7 \text{ c.g.s., or } 260 \text{ gm. mm.}^2.$$

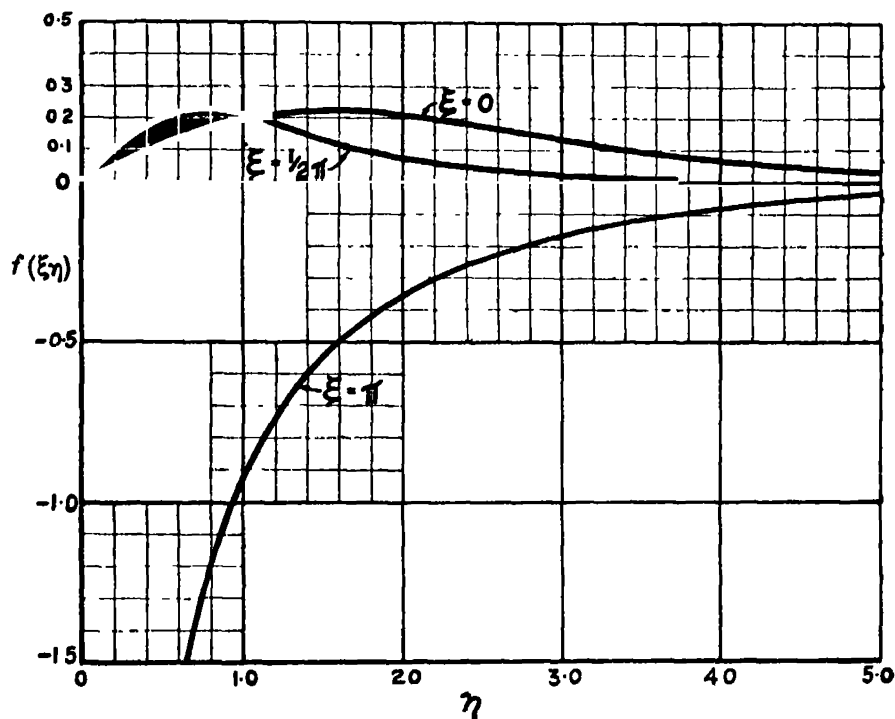


FIG. 6.

For copper  $\mu = 4.5 \times 10^{11}$ , so that  $X_v = 450 \text{ gm. mm.}^2$ .

These values are, in fact, of the same order of magnitude as the observed elastic limits of copper and aluminium. It seems, therefore, that it is justifiable to state that the existence of a finite elastic limit in a crystal is not inconsistent with the hypothesis that centres of dislocation could migrate freely through a *perfect* unstressed crystal of copper or aluminium in a stress-free state (if indeed such a crystal can exist).

**Slip Lines.**—In the preceding pages the distribution of centres of dislocation has been considered as statistically uniform. If, however, a slipping begins in a limited region of the crystal by separation of positive and negative centres of dislocation, the accumulation of positive centres on the right and negative centres on the left borders of this region will produce positive shear stresses

outside the region and negative shear stress inside it. When a positive external shear stress is applied, therefore, the shear stress will be greatest to the right and left of the disturbed region. Thus the application of a positive external stress will tend to extend the disturbed area in the direction of the slip lines. The extension of the disturbed region would thus be analogous to the extension of a crack. It seems, then, that the conception of plastic distortion here put forward leads to an expectation that regions of concentrated slipping would extend in the directions of the slip planes. The existence of slip lines is therefore consistent with the theory though it is inconsistent with the detailed conception of a regular lattice of dislocations which was used in deriving parabolic stress-strain relationship.

### *Summary.*

The fact that the macroscopic distortion of metallic crystals is a shear parallel to a crystal plane and in a crystal direction and the fact that this remains true even when the distortion is large shows that the plastic strain must be chiefly due to the sliding of one plane of atoms over its immediate neighbour in such a way that the perfect crystal structure is re-formed after each atomic jump. It is supposed that slipping occurs over limited lengths  $L$  of the slip plane, and it is shown that this type of plastic strain necessarily gives rise to elastic stresses near the two dislocations which occur at the two ends of each of these lengths  $L$ .

It is then shown that the assumption that such dislocations will migrate through the crystal, owing perhaps to temperature agitation, under the influence of even the smallest shear stress leads to a definite picture of the mechanics of plastic distortion. This theory of strain hardening is expressible in quantitative form and gives a parabolic relationship between stress and plastic strain, namely,  $S/\mu \sqrt{s} = \kappa \sqrt{\lambda/L}$ . This expression is in good agreement with the results of experiment with metals which crystallize in the cubic system.

The observed parabolic relationship is then used in connection with the formula to determine  $L$  which is found, at room temperatures, to be of order of magnitude  $10^{-4}$  cm. This is of the same order of magnitude as the observed spacings of faults in metals and rock salt. According to this theory the part played by the system of faulting or mosaic structure is to limit the free motion of centres of dislocation. The actual strain takes place inside the "blocks" of the mosaic structure and the crystallographic nature of the faults, *i.e.*, whether they are boundaries of dendrites, a superstructure or merely "pores" is immaterial from the point of view of the theory.

---

*The Strength of Rock Salt.*

By G. I. TAYLOR, F.R.S., Royal Society Yarrow Professor.

(Received February 7, 1934.)

The strength of rock salt has attracted a very large number of workers because the equilibrium of the crystal lattice appears to be well understood. The relationship between atomic forces and elastic constants has been calculated using Born's theory, and has been shown to agree with experiment. The force necessary to tear a crystal into two parts across a crystal plane has also been calculated using the same assumptions about atomic forces as those which account for the elastic properties. Very great disagreement, however, was found. Thus Zwicky\* found that a rock salt crystal ought theoretically to stand a stress of 200,000 gm. mm.<sup>2</sup> before breaking while the breaking stress actually observed† in samples tested at "room" temperatures is only about 450 gm. mm.<sup>2</sup>.

The reason for this discrepancy has been discussed by Joffé and others who concluded that it was due to surface cracks. By pulling rock salt crystals at high temperatures Joffé found that they stretch plastically, and that strength increases with the amount of plastic stretching. He gives a curve in which the diminution in area of cross-section of a specimen is related to the increase in strength. With large amounts of stretching the strength could be raised to 5000 gm. per square millimetre. Thus the strength is increased by plastic distortion just as it is for a metal.

Recently Piatti‡ has described experiments in which considerably higher strength has been attained in rock salt crystals which are stretched under controlled conditions in a bath of salt solution at room temperature. These strengths were associated with plastic extension, but the amount of plastic extension necessary to raise the strength to any given amount was very much less than that found by Joffé; thus Joffé gives the extension necessary to attain a strength of 5000 gm. mm.<sup>2</sup> as that which reduces the area of cross-section to 1/30th of its original thickness. This corresponds with an elongation of 3000%. In Piatti's experiments the crystals attained strengths from 10,000 to 20,000 gm. mm.<sup>2</sup> with extension of about 10 to 20%.

\* 'Phys. Z.', vol. 25, p. 223 (1924).

† Joffé, 'Z. Physik,' vol. 22, p. 286 (1924).

‡ 'Z. Physik,' vol. 77, p. 401 (1932); 'Nuovo Cimento,' p. 102 (1932).

When a crystal of rock salt suffers plastic distortion the perfection of the crystal lattice is destroyed so that experiments designed to raise the strength of a rock salt crystal by plastic distortion to the theoretical value calculated for a perfect crystal can only be successful when the crystal is no longer in the condition for which the calculation was made.

For this reason the theoretical interpretations which have hitherto been given concerning the results of experiments on the strength of rock salt seem to be misleading. For a pure and well-annealed rock salt crystal (natural or artificial) the principal factors which determine its strength seem to be the temperature and the amount of plastic distortion, which it has undergone since it was in its undistorted condition. Any theoretical discussion of strength must necessarily involve both these factors. It appears, however, that no one has yet introduced the amount of distortion into theoretical discussion, and it is for this reason that I put forward in the following pages an analysis of experiments on the plasticity of rock salt which is based on a recent theory of the strain-hardening of metallic crystals.\*

This theory is concerned with metallic crystals in which the slipping is parallel to one slip plane. The slipping of a plane of atoms over the adjacent parallel plane is conceived as a jump which would leave the regularity of the crystal structure unaltered if slipping took place over the whole area of the slip plane at once. The characteristic feature of the theory, however, is that the areas on the slip plane over which the jump has occurred are considered to spread out from centres and to be propagated along the plane till they are stopped by some fault or irregularity in the crystal structure. The edges of the areas over which slipping has taken place are called dislocations and the slipping process is regarded as being determined entirely by the motion of these dislocations. The plastic properties of the crystal are determined by the distance  $L$  through which the dislocation can move before its progress is stopped by a fault in the crystal structure.

If  $S$  is the component of shear stress parallel to the slip plane and  $s$  is the total plastic shearing strain since the crystal was in a completely undeformed state the theoretical relationship between  $S$  and  $s$  to which this theory leads is

$$\frac{S}{\mu\sqrt{s}} = \kappa \sqrt{\frac{\lambda}{L}}, \quad (1)$$

where  $\mu$  is the coefficient of elastic rigidity and  $\lambda$  is the atomic spacing along the direction of slip, i.e., the least jump which will leave the crystal structure

\* "The Mechanism of Plastic Deformation of Crystals."



unaltered.  $\kappa$  is a constant which depends on the arrangement of centres of dislocation in the crystal. Values  $\kappa = 0.12$  and  $\kappa = 0.17$  were found for two cubic arrangements.

In applying this theory to rock salt crystals the first step is to find the connection between the plastic stress-strain relations in a tensile test of a crystal of rock salt and those which would be found if the crystal could be caused to slip parallel to one slip plane. It has frequently been remarked that when rock salt is stretched in the plastic condition the distortion does not appear to be due to slipping parallel to one plane. On the other hand, the distortion is definitely related to the crystal structure because in Joffé's experiments a round specimen is described as thinning down to a narrow band of elliptic cross-section. At first sight this seems to show a difference between the behaviour of metals and of rock salt. Further consideration, however, reveals the fact that if rock salt behaves *exactly* like a metal, slip occurring on that slip plane for which the component of shear stress in the direction of shear is greatest, slipping would always be possible on two planes simultaneously instead of on one only. To prove this we may notice that the slip planes are of type (110), and the direction of slip is of type [110]. The six possible slip planes occur in three pairs which intersect on the three cubic axes of the crystal. Considering one such pair we may denote  $\theta_1$  and  $\theta_2$  as the angles between the normals to the planes and the direction of the axis of the specimen. If  $\phi_1$  and  $\phi_2$  are the directions which the projections of the longitudinal axis of the specimen on these planes make with the direction of slip, the component of shear stress parallel to the direction of slip is  $S_1 = P \cos \theta_1 \sin \theta_1 \cos \phi_1$  in one case and  $S_2 = P \cos \theta_2 \sin \theta_2 \cos \phi_2$  in the other. It is easy to prove that when the two slip planes are perpendicular to one another and the direction of slip perpendicular to their line of intersection

$$\cos \theta_1 \sin \theta_1 \cos \phi_1 = \cos \theta_2 \sin \theta_2 \cos \phi_2$$

so that  $S_1 = S_2$ . Hence the components of shear stress on the two members of a pair of slip planes which intersect on a cubic axis are equal to one another. If therefore rock salt obeys the same law of distortion that has been found for aluminium, slipping on two planes simultaneously is always to be expected instead of the single slipping which occurs in general with metals of the cubic system.

The relationship between the extension of a specimen and the amount of slip has been discussed both for double and single slipping. For example, when the axis of the specimen is one of the cubic axes (as it was, for instance, in

Piatti's experiments) we may suppose that during an extension of the specimen from  $l$  to  $l + \delta l$  an amount of slip  $\frac{1}{4}\delta s$  has occurred on each of the four slip planes which are at  $45^\circ$  to the axis. Then  $\delta l = \frac{1}{2}\delta s$  so that  $s = \frac{1}{2} \log(l/l_0)$ , where  $l_0$  is the original length of the specimen. This relationship is still true if only two of the four slip planes considered are operative. The tensile stress  $P$  is then equal to  $2S$ .

Hence

$$\frac{P}{\sqrt{\log(l/l_0)}} = 2\sqrt{2} \frac{S}{\sqrt{s}},$$

so that (1) becomes

$$\frac{P}{\sqrt{\log(l/l_0)}} = 2\sqrt{2} \kappa \mu \sqrt{\frac{\lambda}{r}}. \quad (2)$$

### *Comparison with Observations.*

To compare this formula for the strength of rock salt with observation it is only possible to use data for crystals in which both  $P$  and  $l/l_0$  have been measured. Usually the extension has been measured, if at all, only at the breaking point, so that it is only possible to compare one stage in the process of plastic distortion of one specimen with another stage of the process in that of another specimen. The available data from Piatti's experiments are given in Table I. In that table  $P$  is given in grammes per square millimetre in column 1. In column 2 the value of  $l/l_0$  and in column 3 the value of  $P/\sqrt{\log(l/l_0)}$  also expressed in grammes per square millimetre. The table is arranged so that the figures in column 3 decrease from top to bottom of the column. It will be seen that different specimens vary very largely, the values of  $P/\sqrt{\log(l/l_0)}$  varying from  $75.0 \times 10^3$  to  $6.5 \times 10^3$  gm. mm.<sup>2</sup>. In Table II similar figures are taken from a paper by Smekal.\* In this case the variation in  $P/\sqrt{\log(l/l_0)}$  is not so large as in Piatti's experiments, but the lower limit is nearly the same, namely,  $6 \times 10^3$  gm. mm.<sup>2</sup>.

The great variability in the results given in Tables I and II may be due to the presence of impurity for it has been shown that the introduction of small amounts of foreign material,† e.g., lead chloride and cuprous chloride into a rock salt crystal has a very great hardening effect. It may also be due to varying amounts of distortion or work-hardening in the natural crystals before starting the experiments.

\* 'Phys. Z.,' vol. 32, p. 187 (1931).

† Metag, 'Z. Physik,' vol. 78, p. 363 (1932).

In a more recent set of measurements Thiele\* has obtained the load-extension or strain-hardening curves for several specimens of rock salt at a number of different temperatures from 20° C. to 600° C. Thiele's results also showed

Table I.—Piatti's Measurements.

P.	$l/l_0$ .	$P/\sqrt{\log (l/l_0)}$ .
gm. mm. <sup>2</sup>		gm. mm. <sup>2</sup>
23150	1.100	$75.0 \times 10^3$
19530	1.215	$43.1 \times 10^3$
10775	1.093	$36.0 \times 10^3$
13890	1.187	$33.5 \times 10^3$
9540	1.085	$33.5 \times 10^3$
10290	1.120	$31.0 \times 10^3$
13370	1.215	$30.0 \times 10^3$
12650	1.215	$28.5 \times 10^3$
9920	1.128	$28.0 \times 10^3$
10685	1.190	$26.0 \times 10^3$
10160	1.207	$23.5 \times 10^3$
9260	1.270	$18.9 \times 10^3$
6945	1.155	$18.3 \times 10^3$
8960	1.320	$16.8 \times 10^3$
6250	1.240	$13.4 \times 10^3$
4720	1.140	$13.1 \times 10^3$
5785	1.228	$12.8 \times 10^3$
5340	1.200	$12.8 \times 10^3$
3654	1.097	$12.1 \times 10^3$
3150	1.122	$10.5 \times 10^3$
3470	1.129	$9.7 \times 10^3$
2045	1.102	$6.5 \times 10^3$

Table II.—Measurements given by Smekal, 'Phys. Z.,' vol. 32, p. 187, (1931).

P.	$l/l_0$ .	$P/\sqrt{\log (l/l_0)}$ .
gm. mm. <sup>2</sup>		gm. mm. <sup>2</sup>
4350	1.058	$18.0 \times 10^3$
2660	1.032	$15.0 \times 10^3$
3390	1.051	$15.0 \times 10^3$
2970	1.074	$10.9 \times 10^3$
3430	1.114	$10.2 \times 10^3$
2180	1.103	$6.8 \times 10^3$

considerable variations, but by comparing the load-extension curves of different specimens he was able to distinguish between them in regard to their state of hardness in their natural condition. The yield point of one of his specimens (No. 13), for instance, was about 10 times as high as that of some others while

\* 'Z. Physik.,' vol. 75, p. 763 (1932).

the breaking stress was about seven times as high. By prolonged heating at 600° C. he was able to reduce this particular specimen to a condition similar to that of others. In this way it was found possible to avoid the very great variability among individual measurements which characterized the results of Piatti, and to approximate to a definite standard characteristic of pure rock salt in a work-free\* condition.

The complete results for one such specimen are shown in fig. 1 which is reproduced from Thiele's paper. In this diagram the maximum value of the extension was only 3.5% so that ordinates which are expressed as extension

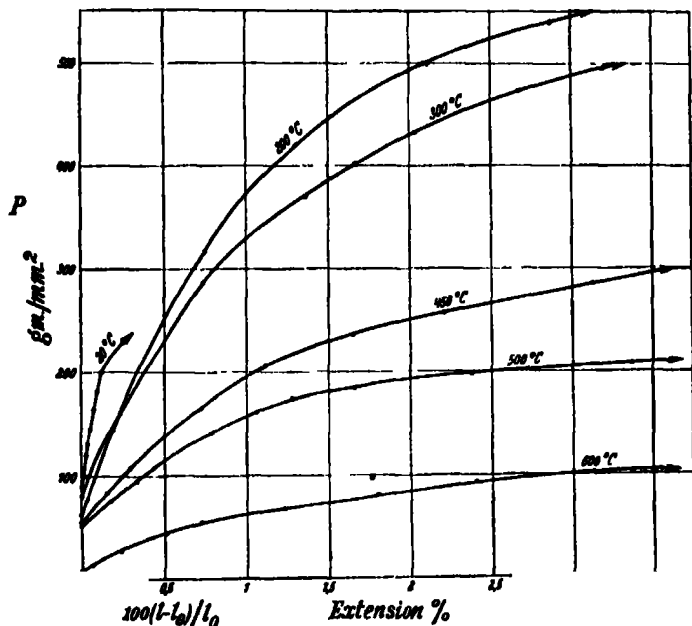


FIG. 1.—Load-extension curves for rock salt after prolonged initial annealing (Thiele).

per cent. are very nearly proportional to  $\log(l/l_0)$ . According to the formula (2) if  $L$  depends only on the temperature and the disposition of pores, system of faulting or superstructure in the crystal, then  $P/\sqrt{\log(l/l_0)}$  should be constant. The theoretical prediction is therefore that the curves shown in fig. 1 should be parabolas.

To test this the points on Thiele's curves representing his actual observations were marked on a diagram, fig. 2, and parabolas drawn so as to pass as nearly as possible through the observed points. It will be seen that the agreement with theory is very good.

\* "Verformungsfrei."

Thiele's results are tabulated in Table III. Values of  $P$  and  $l/l_0$  taken for his curves are given in columns 1 and 2 while in column 3 are given the values of  $P/\sqrt{\log l/l_0}$  calculated from the figures of columns 1 and 2.

It will be seen that at each temperature  $P/\sqrt{\log (l/l_0)}$  is nearly constant. The average value of the constant for each observed temperature is also given in column 3.

It is worth noticing that at  $20^\circ \text{C}$ . the constant is  $6 \times 10^3 \text{ gm. mm.}^2$ , which is identical with the lowest values obtained both in Piatti's and Smekal's experiments. This fact may be significant since Piatti's and Smekal's experiments

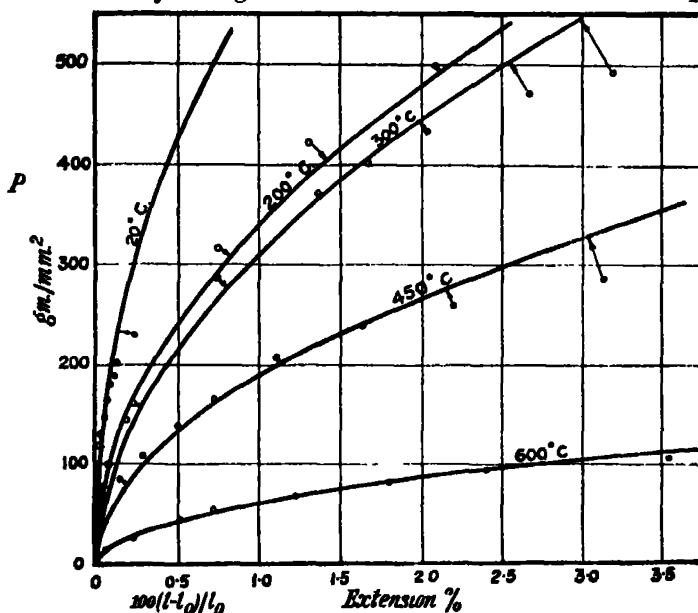


FIG. 2.—Comparison of theory with experiment. The curves represent the theoretical relationship  $P/\sqrt{\log l/l_0} = \text{constant}$ .

were carried out under water which allowed an extension of 10% to take place. Thiele's experiments were not performed under water and the specimen seems to have stretched uniformly only up to about 0.2% extension. This seems to suggest that the presence of water round the specimen may have no effect on the strain-hardening curve though it has a very large effect on the total extension which the specimen can suffer before breaking.

#### Determination of $L$ .

The constancy of the observed values of  $P/\sqrt{\log (l/l_0)}$  at any given temperature seems to confirm the accuracy of the conception here put forward of the

Table III.—Thiele's Measurements (taken from his curves).

P.	$l/l_0$ .	$P/\sqrt{\log (l/l_0)}$ .
Temperature 20° C.		
gm. mm. <sup>2</sup>		gm. mm. <sup>2</sup>
145	1.0005	$6.5 \times 10^3$
165	1.0007	$6.2 \times 10^3$
180	1.0010	$5.7 \times 10^3$
200	1.0013	$5.5 \times 10^3$
	Mean .	$6.0 \times 10^3$
Temperature 200° C.		
140	1.002	$3.1 \times 10^3$
320	1.0075	$3.7 \times 10^3$
420	1.013	$3.7 \times 10^3$
500	1.021	$3.4 \times 10^3$
540	1.029	$3.2 \times 10^3$
	Mean .	$3.4 \times 10^3$
Temperature 300° C.		
160	1.0028	$3.0 \times 10^3$
290	1.0075	$3.3 \times 10^3$
372	1.0140	$3.1 \times 10^3$
400	1.0167	$3.1 \times 10^3$
428	1.0205	$3.0 \times 10^3$
475	1.0269	$2.9 \times 10^3$
495	1.0320	$2.8 \times 10^3$
	Mean	$3.0 \times 10^3$
Temperature 450° C.		
85	1.0015	$2.2 \times 10^3$
106	1.0029	$2.0 \times 10^3$
139	1.0050	$2.0 \times 10^3$
167	1.0073	$2.0 \times 10^3$
206	1.0115	$1.9 \times 10^3$
236	1.0165	$1.8 \times 10^3$
256	1.0221	$1.7 \times 10^3$
285	1.0313	$1.6 \times 10^3$
	Mean .....	$1.9 \times 10^3$
Temperature 600° C.		
15	1.0006	$0.61 \times 10^3$
29	1.0024	$0.59 \times 10^3$
44	1.0051	$0.62 \times 10^3$
55	1.0072	$0.65 \times 10^3$
68	1.0122	$0.62 \times 10^3$
80	1.0180	$0.60 \times 10^3$
92	1.0240	$0.59 \times 10^3$
104	1.0355	$0.55 \times 10^3$
	Mean ... ..	$0.60 \times 10^3$

mechanism of plastic deformation. Accordingly we may proceed with more confidence to use (2) for the determination of  $L$ . This formula may be re-written

$$L = 8\mu^2\kappa^2\lambda [P/\sqrt{\log(l/l_0)}]^{-2}. \quad (3)$$

For rock salt  $\mu$  may be taken as  $1.3 \times 10^{11}$  c.g.s.  $\lambda$  is the least relative distance through which two parts of a crystal situated on opposite sides of a slip plane must move in order that the perfect crystal structure may be re-formed. For the cubic lattice of NaCl  $\lambda = \frac{1}{2}\sqrt{2} \times (\text{side of unit cube}) = \frac{1}{2}\sqrt{2} (5.628 \times 10^{-8}) = 3.98 \times 10^{-8}$  cm.

The value of  $\kappa$  depends on the manner in which the dislocations are distributed. Values  $\kappa = 0.12$  and  $0.17$  were calculated for special cases (i.e., square lattices of dislocations). In the absence of any definite method for estimating  $\kappa$  it is only possible to point out that the various assumptions regarding the distribution of centres of dislocation lead to values which do not differ in order of magnitude and to calculate the values of  $L$  which correspond with the various assumptions.

Setting  $\kappa = 0.12$  and  $0.17$  in (2) and using the observed values of  $P/\sqrt{\log(l/l_0)}$  the two sets of values given in Table IV have been calculated for  $L$ .

Table IV. —Values of  $L$  calculated from formula

$$L = 8\mu^2\kappa^2\lambda [P/\sqrt{\log(l/l_0)}]^{-2}.$$

Temperature.	$P/\sqrt{\log(l/l_0)}$ .	$L$ (cm.).	
		$\kappa = 0.12.$	$\kappa = 0.17.$
° C.	gm. mm. <sup>2</sup>		
20	$6.0 \times 10^3$	$2.2 \times 10^{-4}$	$4.4 \times 10^{-4}$
200	$3.4 \times 10^3$	$6.8 \times 10^{-4}$	$1.4 \times 10^{-3}$
300	$3.0 \times 10^3$	$8.8 \times 10^{-4}$	$1.9 \times 10^{-3}$
450	$1.9 \times 10^3$	$2.2 \times 10^{-3}$	$4.4 \times 10^{-3}$
600	$0.6 \times 10^3$	$2.2 \times 10^{-3}$	$4.4 \times 10^{-3}$

*Comparison between  $L$  and the Average Size of an Element of a Mosaic Crystal Structure.*

It will be seen from Table IV that at 20° C.  $L$  is of order of magnitude  $10^{-4}$  cm. This is of the same order of magnitude as the estimates which have already been made of the average distance apart of pores or faults in the

structure of rock salt. From X-ray data Darwin, Bragg, and James\* concluded that the linear dimensions of the elements of their mosaic are greater than  $0.5 \times 10^{-4}$  cm. From the work of Seidentopf† on the diffusion of the vapour of alkali metals into rock salt Smekal‡ considered that the average distance apart of "Löckerstellen" or loose places in the structure is about  $2 \times 10^{-4}$  cm.

It seems likely, therefore, that the "löckerstellen" of Smekal or the faults at the boundaries of the blocks of a mosaic structure do, in fact, act as barriers to the free movement of dislocations. If the present theory is accepted these faults control the plastic properties of the crystal by limiting the movement of dislocations, but the actual slipping occurs in the parts of the crystal where the atoms lie in perfect order. This conception of the mechanics of plastic deformation is the inverse of what has sometimes been proposed, namely, that the slipping occurs in the faults leaving the intervening blocks as fragments in perfect crystalline order.

At high temperatures the mean free path  $L$  of the dislocations is many times as great as the mean size of an element of the mosaic. This might be accounted for by supposing that dislocations can cross the faults or boundaries between the elements of a mosaic at some places and not at others. At high temperatures a fault would act as a barrier only at points where the discontinuity between the crystal structure and neighbouring blocks was a maximum. At lower temperatures a greater proportion of the area of the fault surfaces would be opaque to dislocations. This suggestion has been partially analysed in a previous paper in connection with aluminium.

### *The Elastic Limit.*

In the previous paper a suggestion was made that the elastic limit was caused by internal elastic stresses produced by the misfit at the boundaries of elements of a mosaic. This idea led to an approximate formula for determining the order of magnitude of the elastic limit. When the average angle between the orientation of the crystal axes of neighbouring blocks is 1 minute of arc and the average diameter of a block is  $10^{-4}$  cm. the value found for the elastic limit was of order of magnitude  $10^{-4} \mu$ . For rock salt  $\mu = 1.3 \times 10^{11}$ , so that the order of magnitude of the elastic limit which would result from a block structure of this type would be  $1.3 \times 10^7$  c.g.s. or 130 gm. mm.<sup>2</sup>.

\* 'Phil. Mag.' vol. 1, p. 897 (1926).

† 'Phys. Z.', vol. 6, p. 855 (1905).

‡ 'Phys. Z.', vol. 26, p. 709 (1925).



Referring to fig. 1 it will be seen that the elastic limit of the specimen there referred to was about 90 gm. mm.<sup>2</sup>, at 20° C., and that it decreased to 50 gm. mm.<sup>2</sup> at 500° C., and 10 gm. mm.<sup>2</sup> at 600° C. The theory, therefore, gives at any rate the correct order of magnitude for the elastic limit.

*Summary.*

The strength of rock salt has so far only been studied from the theoretical point of view without considering the effect of plastic strain. Experiment shows, however, that plastic strain is the main factor determining the strength of well-annealed crystals. A recent theory of the strength of metals is applied to rock salt and shown to lead to a parabolic relationship between tensile stress,  $P$ , and plastic strain  $\log (l/l_0)$  where  $l$  is the length of a stretched bar of rock salt whose initial length is  $l_0$ . This relationship is

$$P/\mu \sqrt{\log (l/l_0)} = 2 \sqrt{2} \kappa \sqrt{\lambda/L},$$

where  $\mu$  is the elastic modulus of rigidity,  $\lambda$  is the atomic spacing along the plane of slip,  $\kappa$  is a constant whose value can be calculated approximately,  $L$  is the mean free path of centres of dislocation.

Comparing this with observations of plasticity in rock salt good agreement is found, and  $L$  is shown to be of order of magnitude  $10^{-4}$  cm. This is the order of magnitude of the distance which has been observed between faults in the structure of rock salt. It is concluded that the strain in rock salt occurs in the crystalline parts of the structure where the crystal order is perfect, and that the strength is determined by the mean free path  $L$  of the centres of dislocation.  $L$  is determined by the distance apart of the faults and by the temperature. The theory therefore assigns a definite function to the faults in determining the strength of crystals irrespective of their actual crystallographic or atomic nature.

---

*The Vertical Distribution of Ozone in the Atmosphere.*

By F. W. PAUL GÖTZ, Dr. Phil. Nat. (Lichtklimatisches Observatorium, Arosa, and University of Zürich), A. R. MEETHAM, D.Phil., and G. M. B. DOBSON, D.Sc., F.R.S. (Reader in Meteorology, University of Oxford).

(Received February 16, 1934.)

*List of Symbols used in the Paper.*

$I$  = the intensity of radiation of wave-length 3110 Å. received by instrument.\*

$I'$  = the intensity of radiation of wave-length 3290 Å. received by instrument.

$\alpha$  = the absorption coefficient per centimetre of ozone for  $\lambda = 3110$  Å.

$\alpha'$  = the absorption coefficient per centimetre of ozone for  $\lambda = 3290$  Å.

$\beta$  = the scattering coefficient of the whole atmosphere for  $\lambda = 3110$  Å.

$\beta'$  = the scattering coefficient of the whole atmosphere for  $\lambda = 3290$  Å.

$I_0$  = the intensity of 3110 Å. which would be received if there were no atmosphere.

$I'_0$  = the intensity of 3290 Å. which would be received if there were no atmosphere.

$Z$  = the sun's zenith distance as seen from the ground.

$\zeta_h$  = the sun's zenith distance as seen from a point on a line joining the instrument and the sun, and at a height  $h$  above the ground. Thus

$$\sin \zeta_h = R \sin Z / (R + h),$$

where  $R$  is the earth's radius.

$x$  = the total amount of ozone, in centimetres at N.T.P.

$x_1, x_2, x_r$ , etc = the amount of ozone in layers 1, 2,  $r$ , etc.

$\beta_r = \beta \times (\text{Mass of air in layer } r) / \text{Total mass of atmosphere.}$

\*  $I$  and  $I'$  may refer either to direct sunlight or to light from the zenith sky, with corresponding different values.

$$L = \frac{(\text{The path length of the solar beam through the atmosphere})}{(\text{Path length for vertical sun and pressure} = 760 \text{ mm.})}.$$

$l$  = the path length in centimetres of ozone of the solar beam.

$b$  = the atmospheric pressure at the height of scattering (in millimetres Hg).

$\zeta_{rs}$  = the value of  $\zeta_h$  when  $h$  is the height of the absorbing layer ( $s$ ) above the scattering layer ( $r$ ).

## PART I.—METHODS OF OBSERVATION.

### I. *Introduction.*

Several estimates of the average height of the ozone in the atmosphere have been made in recent years. In most cases the spectrum of direct sunlight has been measured as the sun was rising or setting, and by comparing the relative intensities of two wave-lengths—one being strongly absorbed by the ozone and the other little absorbed—the total amount of ozone traversed by the light in its oblique path through the atmosphere, was determined. Assuming the amount of ozone in the atmosphere to remain constant during the observations, and to be concentrated in a thin layer, the height of the ozone was found in terms of the radius of the earth. This method is not very satisfactory since the accuracy is poor except when the sun is very low. In the former measurements photographic methods were used, and except in high latitudes it was not possible to make measurements when the sun's zenith distance was greater than about  $77^\circ$  since the intensity of the shorter wave-lengths was too weak. Further, certain constants of the instruments, which are difficult to measure accurately, greatly affect the height deduced.

The improved means now available for measuring the amount of ozone in the atmosphere, using photoelectric cells,\* have made possible an entirely different method of estimating the height. The new instruments measure, as before, the relative intensities of two wave-lengths which are absorbed to very different extents by ozone (3110 Å. and 3290 Å.), but the greatly increased sensitivity of the instrument enables measurements to be made either on the direct light from the sun or on the light from the clear blue zenith sky. In the latter case it is possible to continue the observations till well after sunset or to start before sunrise, and such observations on the light from the zenith sky can be used to deduce the average height of the ozone with considerable

\* 'Proc. Phys. Soc. Lond.,' vol. 43, p. 324 (1931).

accuracy, and also to obtain an idea of the way in which it is distributed, vertically in the atmosphere.

Observations on the light of the zenith blue sky differ from those on direct sunlight; in the latter, the light must necessarily traverse an amount of ozone equal to that in a vertical column of the atmosphere multiplied by the secant of the sun's zenith distance as seen from the ozone layer. The amount of ozone which the light must traverse continually increases as the sun's zenith distance increases. Light received from the clear blue zenith sky is quite different, since light scattered downwards from the direct solar beam at all heights in the atmosphere will contribute to the light received by the instrument. The amount of light reaching the instrument which was scattered from the direct solar beam at any particular height will depend on two things: (1) the amount of air at that height available to scatter the light; (2) the absorption of the light before and after scattering. As the height of scattering is increased the effect of (1) will be to decrease the light received, while the effect of (2) will be to increase the light received, since the oblique path through the atmosphere is replaced more and more by the shorter vertical path. It can be shown that there will be a limited region where most of the light actually received has been scattered, and we may regard the height of this region as the effective height of scattering. This effective height of scattering will depend on the rate of absorption of the light and on the zenith distance of the sun. It will be greater for those wave-lengths which are more rapidly absorbed, and will also be greater when the sun is low than when it is high. Thus it comes about that 3110 Å.—being more rapidly absorbed than 3290 Å.—will always be scattered at a greater effective height, and when the sun is only a few degrees above the horizon the effective height of scattering for 3110 Å. will be above most of the ozone, while 3290 Å. is scattered much lower down. After this the intensity of 3110 Å. will change but slowly as the sun sets, while 3290 Å. is still rapidly decreasing in intensity. Hence the ratio of 3110 Å. to 3290 Å. which decreases at first as  $Z$  increases, reaches a minimum (at  $Z = 85^\circ$  approx.) and increases again for still larger values of  $Z$ .

In the limiting case when all the light is scattered so high in the atmosphere that there is negligible absorption before scattering takes place, all the absorption will occur on the vertical part of the path and the ratio of the two wave-lengths will be approximately the same as that when the sun is nearly overhead. It must be understood that it is only the ratio of the intensities of the two wave-lengths which increases when the sun is low and that the absolute intensity of all wave-lengths continuously decreases as the sun sets.

The fact that the sky-light is relatively richer in the shorter wave-lengths when the sun is setting than when it is a few degrees up, was first noticed by Götz in 1929 during observations at Spitzbergen, and called by him the "Umkehr Effect."\* He also called attention to the possibility of deducing the vertical distribution of the ozone in the atmosphere from such observations of the zenith blue sky, and the present work is an attempt to obtain accurate data for all heights of the sun, for all seasons of the year and for different amounts of ozone, and to deduce the distribution of the ozone from these observations. Troubles due to an unknown amount of scattering by dust and haze have been largely avoided by taking the observations in the clear air at Arosa, Switzerland, at a height of 1860 metres. In 1932 two photoelectric instruments were taken to Arosa, and for five weeks simultaneous observations were made both on direct sunlight and on the light from the clear blue zenith sky on all clear days from before sunrise to after sunset. Unfortunately bad weather prevented many days' observations being obtained. Since then observations have been continued with a single instrument at Arosa, but with only one instrument it is not possible to take observations sufficiently rapidly when the sun is low on both direct sunlight and zenith skylight; thus one type of observation (usually direct sunlight) had to be omitted for low sun.

The quantity measured by the instrument is the ratio of the intensity of the wave-length 3110 Å. to that of 3290 Å. multiplied by an unknown constant, and it is conveniently expressed as the logarithm ( $-\log I/I' + C$ ),† where  $I$  is the intensity of 3110 Å.,  $I'$  that of 3290 Å. and  $C$  a constant. The results of a day's observations are conveniently shown by a plot of  $-\log I/I'$  against  $Z$ , the sun's zenith distance; or better still against some higher power of  $Z$  since it is desirable to open out the scale near sunrise or sunset when the changes take place rapidly. Plots of  $-\log I/I'$  against  $Z^4$  have been found most generally convenient. Such curves will be referred to as "Umkehr Curves" since they show the "Umkehr Effect" when the sun is low. All the observations on each day have been plotted in this manner and a smooth curve drawn through the points.

On each day one or more observations have been taken on direct sunlight and from these the total amount of ozone has been deduced in the usual way.

In fig. 1 the values of  $-\log I/I'$  are plotted against  $Z^4$  for five days on which the total ozone content ( $x$ ) was nearly identical. It will be seen that all the

\* Götz, 'Gerlands Beitr.,' vol. 31, p. 119 (1931).

† The unknown constant is always eliminated in the calculations and is therefore of no importance in this work and is omitted in the remainder of the paper.

points lie on a narrow band, the scatter of the points when  $Z$  is large being due to the larger error of the observations when the light is weak. This is generally found to be true of any group of days on which the total ozone content is nearly identical and we may conclude that the shape of the "Umkehr Curve" depends mainly on the value of  $x$ .

It is convenient also to plot the results in a different way. In fig. 2 the values of  $-\log I/I'$  for certain fixed values of  $Z$  are plotted against the total

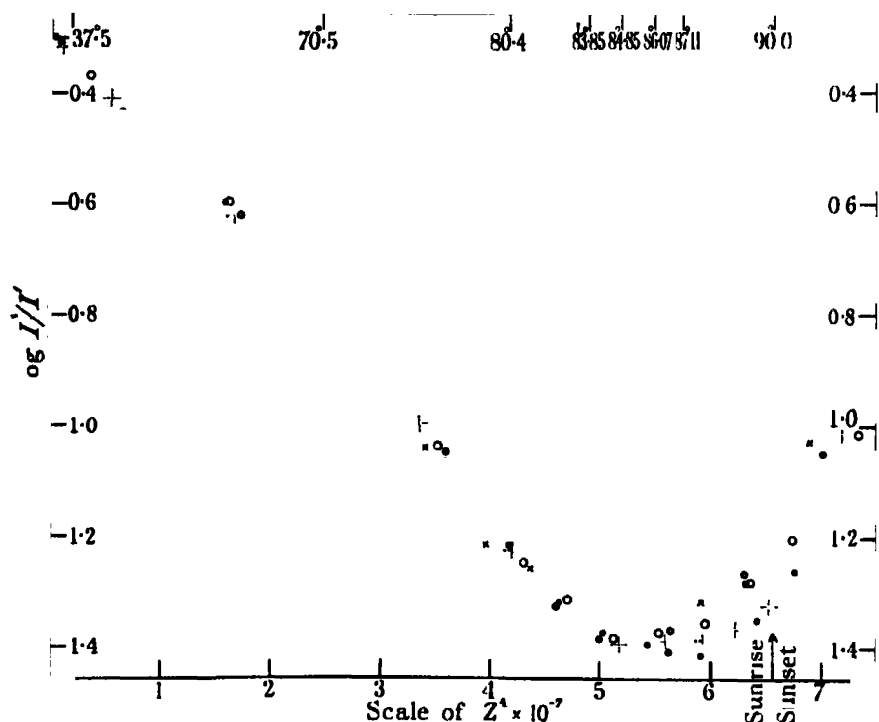


FIG. 1.—Observed values of  $-\log I/I'$  for zenith blue sky on 5 days on which the total ozone content was nearly identical. O, February 14, 1933, a.m. ( $x = 253$ ); ×, February 17, 1933, a.m. ( $x = 260$ ); ●, May 22, 1933, a.m. ( $x = 260$ ); +, June 3, 1933, a.m. ( $x = 263$ ); ○, June 3, 1933, p.m. ( $x = 256$ ).

ozone content ( $x$ ). Again the points lie on narrow bands, the error of observation making the bands wider for the larger values of  $Z$ . A mean line can be drawn through these bands, which, since it depends on a large number of observations, will have a smaller casual error than individual readings. Average "Umkehr Curves" can be then drawn from these mean lines for any value of  $x$ . Such curves, based on all the (46) observations during the year 1932-33, form the most reliable set of data which we have, and are used in the following

calculations, Table I. If the observations obtained in different seasons of the year are plotted separately in the form of fig. 2 it is found that all seasons give the same curve, again showing that the shape of the "Umkehr Curve,"

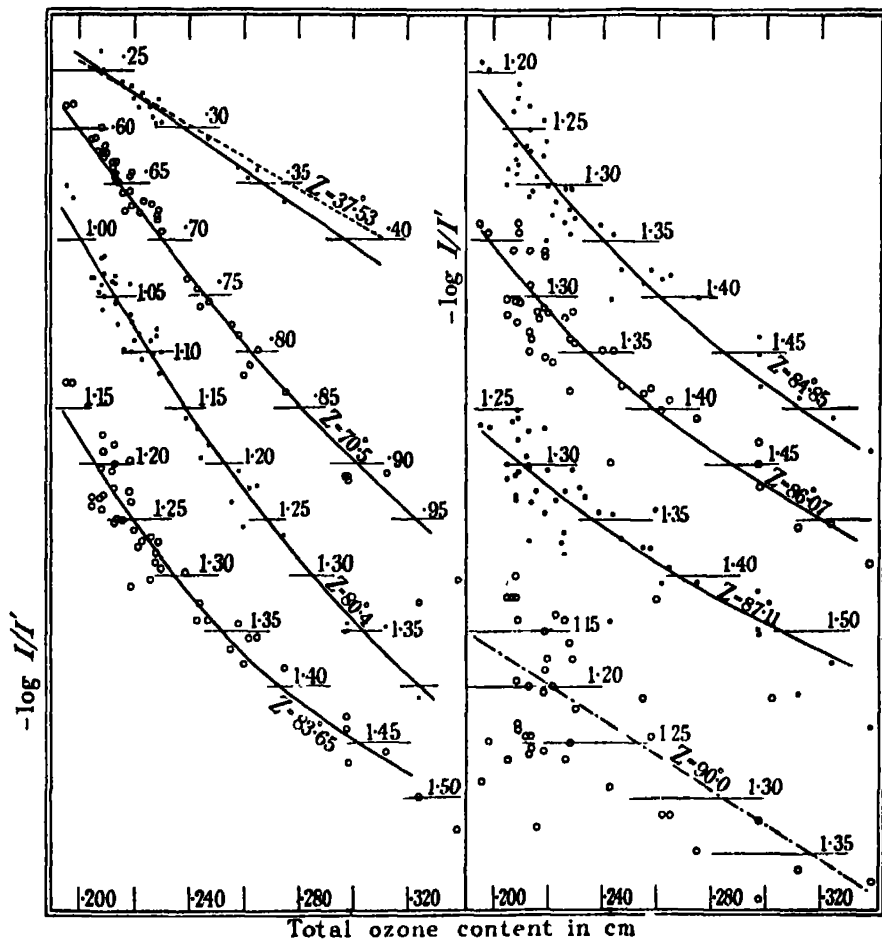


FIG. 2.—“ $-\log I/I'$ ,  $x$ ” curves employing all observations between June, 1932, and June, 1933. Discrete points are taken from the 46 individual days' smooth “Umkehr Curves.” The continuous line for  $Z = 37.53^\circ$  is the best straight line through the discrete points; the dotted curve is the calculated line, neglecting secondary scattering. The broken line for  $Z = 90^\circ$  is the best straight line through the discrete points. The smooth curves for all other values of  $Z$  are calculated by method B, for the distributions given in Table X.

and therefore the vertical distribution of ozone, depends mainly on the total ozone content. “Umkehr Curves” obtained in the morning and afternoon of days when the total ozone content is steady appear to be identical.

Table I.—Mean values of  $(-\log I/I' + \text{const.})$  from zenith sky observations for even values of total ozone content. (Obtained from curves of fig. 2.)

Z.*	$10^{-7} \times Z^{\Delta}$	Total ozone content in cm.					
		0.200.	0.220.	0.240.	0.260.	0.280.	0.300.
°							
37.5	0.20	0.237	0.271	0.306	0.339	0.372	[0.406]
70.5	2.47	0.582	0.603	0.737	0.797	0.855	0.913
80.4	4.18	0.985	1.080	1.158	1.228	1.288	1.344
83.65	4.90	1.164	1.258	1.315	1.360	1.403	1.447
84.85	5.19	1.219	1.301	1.350	1.388	1.426	1.463
86.07	5.49	1.262	1.322	1.360	1.396	1.431	1.468
87.11	5.76	1.272	1.315	1.353	1.387	1.418	1.448
90.0	6.56	1.167	1.198	1.230	1.262	1.293	1.325

\* The values of Z which are used have been chosen to give as many points on the "Umkehr Curves" as possible which are reasonably independent of each other.

*Conversion of Instrument Readings to  $-\log I/I'$ .*—The photometric balancing of the two wave-lengths is affected by an optical wedge in the path of the wave-length 3290 Å. This optical wedge must be calibrated as accurately as possible. For this the instrument is arranged so that only the wave-length 3290 Å. reaches the photoelectric cell and the deflection of the galvanometer is then nearly proportional to the intensity of the light falling on the cell. A metal gauze is held in front of the entrance window of the instrument and the galvanometer reading noted. The gauze is then removed and the wedge adjusted by turning the dial until the same galvanometer reading is obtained. The shift of the wedge is then equivalent to the known transmission of the gauze. These measurements are taken all along the length of the wedge and a curve is drawn showing the value of  $-\log I/I'$  for any dial reading (R). For rough work the wedge can be assumed to be uniform and a "wedge constant"  $k$  calculated so that  $-\log I/I' = kR + \text{const.}$ ; but for accurate work the calibration curve must be used.

## II. Errors of Observation.

In calculating the height and distribution of the ozone in the atmosphere the values of  $-\log I/I'$ ,  $x$  and  $Z$  given by the smooth "Umkehr Curves" for even values of  $x$  (Table I) have been used as the observational basis. These figures were derived directly from the " $-\log I/I'$ ,  $x$ " plots of fig. 2, hence it is important to know the standard error of the values obtained.



The scatter of the points in fig. 2 from the mean smooth curves may be due to any of the following reasons:—

- (1) Error of observation of  $-\log I/I'$  in the zenith blue sky measurements.
- (2) (a) Error of observation of  $-\log I/I'$  in the direct sun measurements leading to any error in the calculated value of  $x$ . (b) Variations of  $x$  during the day on days when only one or two direct sun measurements were made, leading us to take the ozone as constant throughout the day whereas really it may have varied.
- (3) Real differences in the vertical distribution of the ozone in the atmosphere on days having the same value of  $x$ .

The magnitude of (1) can be found by plotting the original readings on individual days and finding the scatter of these readings from the mean curve for the day, the process being repeated for a number of days. Since approximately 16 individual observations are used in obtaining one point in fig. 2, the standard error of a point will be about one-quarter that of individual readings.

An estimate of the magnitude of (2) may likewise be made by using those days on which a large number of direct sun observations were obtained and finding the standard deviation of the individual values of  $x$  from the mean for the day. Then, knowing approximately the variation of  $-\log I/I'$  (for zenith sky observations) with  $x$ , we can calculate the total scatter to be expected in the points of fig. 2 due to the errors of observation. If the observed scatter of the points in this figure be greater than that calculated from the errors of observation, it will, presumably, be due to cause (3). It will be noticed that since the zenith blue sky observations at high sun are nearly independent of the distribution of ozone, cause (3) will not affect the scatter of the points at small values of  $Z$ , and the observed and calculated standard deviations should agree in any case. The standard deviations are given in Table II.

It will be seen that the actual standard deviation of the points in fig. 2 is always greater than that to be expected from the estimated error of the observations, though the difference is insignificant when  $Z$  is small. This indicates that the vertical distribution of ozone is not solely dependent on the value of  $x$ , but that on two days having the same value of  $x$  the vertical distribution of the ozone may not be exactly alike.

It is of interest to calculate the error of observation for measurements of direct sunlight in the same way as that of Table II, and Table III gives such results. The standard error due to random causes is very small.

Table II.—Standard errors of zenith sky observations.

Z:	37.5°	70.5°	80.4°	83.6°	84.8°	86.1°	87.1°	90.0°	91.1°
$Z^{\circ} \times 10^{-2}$	0.20	2.47	4.18	4.90	5.19	5.49	5.76	6.56	6.90
(1) Error of individual reading of — $\log I/I'$ for zenith blue sky	0.005	0.007	0.019	0.030	0.036	0.042	0.048	0.066	0.074
(2) Error of smooth "Umkehr Curve" for any one day . . . . .	0.001	0.002	0.005	0.008	0.009	0.010	0.012	0.016	0.018
(3) Error of mean $x$ for one day obtained from direct sun observations*	0.002 cm.								
(4) Error of points of fig. 2 calculated from (2) and (3) above	0.007	0.008	0.009	0.010	0.011	0.012	0.013	0.017	0.019
(5) Actual deviation of points in fig. 2	0.008	0.012	0.015	0.017	0.020	0.023	0.028	0.064	—
(6) Error of mean "Umkehr Curves" for even ozone values used in calculations . . . . .	0.004	0.006	0.007	0.009	0.010	0.012	0.014	0.032	—

\* This figure is different from the error given in Table III because it includes variations in  $x$  over a period greater than a few minutes, whereas Table III does not.

Table III.—Standard error of direct sun observations and resulting error in total ozone content.

Z	0°	60.3°	71.1°	76.3°	79.4°	81.5°	83.1°
$\sec \zeta_0$	1.0	2.0	3.0	4.0	5.0	6.0	7.0
Error of one observation of — $\log I/I'$	0.0015	0.0030	0.0048	0.0069	0.0092	0.0118	0.0146
Error of mean of four ob- servations . . . . .	0.0007	0.0015	0.0024	0.0034	0.0046	0.0059	0.0073
Resulting error in ozone content (cm.) . . . . .	0.00064	0.00068	0.00073	0.00078	0.00083	0.00089	0.00095

### III. Direct Sun Observations. Method of finding the Amount of Ozone, its Average Height and the Value of $\log I_0/I'_0$ .

Assuming the ozone to be concentrated in a thin layer at a height  $h$ , the formula connecting  $\log I/I'$  as obtained from direct sun observations, with the other quantities is\*

$$\log I/I' = \log I_0/I'_0 - (\alpha - \alpha') x \sec \zeta_h - (\beta - \beta') \sec Z, \quad (1)$$

where

$\log I_0/I'_0$  is the value of  $\log I/I'$  which would be obtained by the instrument if there were no atmosphere and no ozone ;

\* In practice the term  $\sec \zeta_h$  becomes appreciably inaccurate when  $Z > 83.5^\circ$ . Very few observations have been taken at such low altitude of the sun.

$\alpha$  and  $\alpha'$  are the absorption coefficients of ozone for the wave-lengths 3110 Å. and 3290 Å. ;

$\beta$  and  $\beta'$  are the scattering coefficients for air for the same wave-lengths ;

$x$  is the amount of ozone in centimetres at N.T.P.

$Z$  is the sun's zenith distance ; and

$\zeta_h$  is the zenith distance of the sun as it would be observed at a point where the rays from the sun to the instrument pass through the centre of the ozone layer.

$h$  is the height of centre of the ozone layer.

The unknown quantities in this equation are  $\log I_0/I'_0$ ,  $x$  and  $\zeta_h$  (or  $h$  since  $\zeta_h$  is a known function of  $h$ ). Given a series of direct sun observations of  $\log I/I'$  on one day in uniform atmospheric conditions, in which  $Z$  varies from,

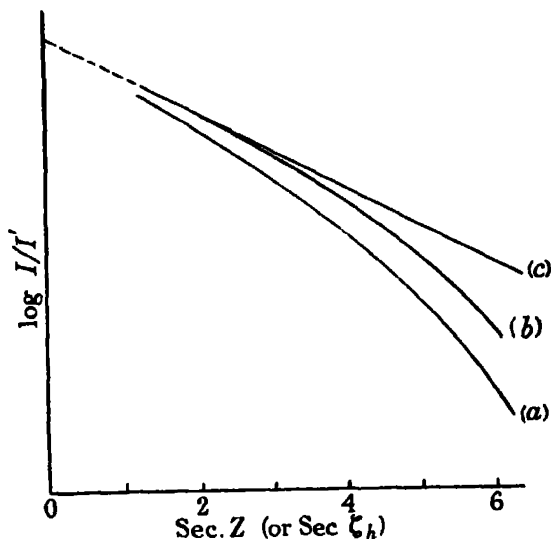


FIG. 3.

say,  $30^\circ$  to  $85^\circ$ , it is possible to evaluate these quantities, but only if they are assumed to remain constant during the observations. It is reasonable to assume that  $\log I_0/I'_0$  which depends on the constancy of the sun and the instrument, does not vary appreciably, but we have no *prima facie* justification for assuming that  $x$  and  $h$  are constant.

In fig. 3 let curve (a) represent the smooth curve obtained from a typical series of observations on direct sunlight,  $\log I/I'$  being plotted against  $\sec Z$  as

abscissa. The method which has been employed for determining the unknowns is as follows :—

- (1) To each ordinate add the corresponding value of  $(\beta - \beta') \sec Z$ ,\* thus forming the curve (b). Then, by equation (1) curve (b) represents the variation of  $\log I_0/I'_0 - x \sec \zeta_h$ .
- (2) If now instead of  $\sec Z$ , we used  $\sec \zeta_h$  as abscissa the curve would be a straight line. We therefore draw a series of curves such as (c) in which the abscissa  $\sec Z$  has been replaced by  $\sec \zeta_h$  corresponding to various assumed values of  $h$  (10 km., 20 km., 30 km., etc.). Whichever of these curves is a straight line determines the true value of  $h$ . If none of these curves is a straight line over the range  $Z = 0^\circ$  to  $Z = 83.5^\circ$  then the assumption that  $x$  and  $h$  are constant is untenable. In practice it is found that, on days with steady meteorological conditions, one of the curves is a straight line within the limits of accuracy of the observations.
- (3) When the curve (c) has been drawn corresponding to the correct value of  $h$ , it is produced to meet the axis of  $\log I/I'$ . Then by equation (1) the point where the line cuts the axis corresponds to  $\log I_0/I'_0$ . In practice values of  $\log I/I'$  at very large values of  $Z$  are not used in this construction.
- (4) Finally  $x$  could be determined from the slope of the line  $c$ , which by equation (1) is  $(\alpha - \alpha') x$ .

In practice while  $\log I_0/I'_0$  and  $h$  are obtained as described above,  $x$  is usually obtained by taking the mean values of  $\log I_0/I'_0$  and  $h$  for a number of good, steady days and assuming them to remain constant. In this way it is possible to obtain an accurate measurement of  $x$  from one observation at any time of day. The very few determinations of  $h$  by direct sun observations which it has been possible to make with the new photoelectric instrument at Arosa indicate a height of about 27 km. above M.S.L. The results obtained with the photoelectric instrument are so much more accurate than those obtained with the older photographic instrument that the lower value given by the new method (27 km.) has probably much more weight than that given previously (50 km.) in spite of the relatively small number of observations which it has been possible to obtain up to the present time with the newer method.

\* It is easily shown that a small error in the assumed value of  $(\beta - \beta')$  causes no appreciable error in what follows.

PART II.—CALCULATION OF THE DISTRIBUTION OF THE OZONE FROM  
ZENITH SKY OBSERVATIONS.

Neglecting for the moment the effect of secondary scattering let us consider the light scattered at any height into the instrument. The intensity of light of any given wave-length reaching any point at a height  $H$  in the atmosphere will be

$$I_0 10^{-\int_H^{\infty} (\alpha x_h + \beta_h) \sec \zeta_h dh},$$

where

$I_0$  is the intensity of radiation received outside the atmosphere ;

$x_h$  is the amount of ozone at any height  $h$  ; and

$\alpha$  is its absorption coefficient ;

$\beta_h$  is the scattering coefficient of the air at a height  $h$  ;

$\zeta_h$  is the zenith distance of the sun as seen from the point in the beam at a height  $h$ .

The amount of light scattered vertically downwards will be

$$K\rho_H (1 + \cos^2 Z) \times I_0 10^{-\int_H^{\infty} (\alpha x_h + \beta_h) \sec \zeta_h dh},$$

where  $K$  is some constant, and  $\rho_H$  is the density of the air at a height  $H$ . The amount of light reaching the instrument which was scattered at a height  $H$  will be

$$K\rho_H (1 + \cos^2 Z) \times I_0 10^{-\int_H^{\infty} (\alpha x_h + \beta_h) \sec \zeta_h dh} \times 10^{-\int_0^H (\alpha x_h + \beta_h) dh}.$$

Hence the amount of light reaching the instrument from all heights, *i.e.*, the total light from the zenith sky will be

$$I = KI_0 (1 + \cos^2 Z) \int_0^{\infty} \rho_h \{10^{-\int_h^{\infty} (\alpha x_h + \beta_h) \sec \zeta_h dh} \times 10^{-\int_0^h (\alpha x_h + \beta_h) dh}\} dh. \quad (1)$$

Since in this work we are always dealing with the ratio of the intensities of two wave-lengths, terms which affect all wave-lengths alike can be neglected : thus  $(1 + \cos^2 Z)$  can be omitted. Further, since it will be necessary to compare the values at any height of the sun with that at some fixed height, all terms independent of  $Z$  and the total amount of ozone, such as the constant  $K$  can also be omitted.

The difficulties of any strict mathematical method of finding the vertical distribution of ozone in the atmosphere from a given "Umkehr Curve" are very great and approximate methods have had to be used. Two different methods have been employed at Arosa and at Oxford in working up the observa-

tions and are described in detail later. The two methods give similar results and are briefly as follows.

*Method A (Arosa).*—The atmosphere is divided into five sections, viz., 2 to 5 km. (the height of Arosa being approximately 2 km.), 5 to 20 km., 20 to 35 km., 35 to 50 km., and above 50 km. The amount of ozone in the highest section is assumed negligible, and that of the lowest three kilometres is taken as 3% of the total ozone in accordance with measurements at ground level. The ozone in the sections 20 to 35 km. and 35 to 50 km. are taken as unknowns, while that in the section 5 to 20 km. is obtained by subtraction from the total ozone content as obtained from direct sun observations, once that in the other sections has been evaluated. The ozone in each section is supposed to be distributed uniformly.

To obtain the two unknown amounts of ozone, the values of  $-\log I/I'$  given by a mean observed "Umkehr Curve" at two values of  $Z$ , are used. It is then possible to deduce equations from which the values of the two unknowns can be exactly evaluated within the accuracy allowed by the observations and the assumptions to be discussed later. Independent solutions can be obtained by using other values of  $Z$  and if desired, dividing the atmosphere into different sections.

*Method B (Oxford).*—In this method the whole atmosphere is divided into eight sections and arbitrary amounts of ozone are assumed to be in each section. By calculating the scattering and absorption of light at all heights under these conditions a theoretical "Umkehr Curve" is calculated. This will, in general, differ from the observed curve and the assumed distribution of ozone in the various sections is varied until a reasonably good agreement is obtained between the calculated and observed curves. In general more than one distribution of ozone can be found which will give a theoretical "Umkehr Curve" which agrees with the observed curve within the limits of error of observation, but it is noteworthy that the average height of the ozone is found to be almost identical for all such distributions.

#### IV. *Method A of Calculating the Vertical Distribution of Ozone.*

(a) *Method.*—The total amount of ozone ( $x$ ) represents the sum of the amounts  $x_r$  in the different sections  $r$  into which the atmosphere is divided. Our aim is to make use of the "Umkehr Curve" for deriving the values  $x_r$ .

If we proceed in a strict mathematical manner the difficulties of a numerical solution of the equations increase rapidly with the number of unknowns.

For this reason we take only two unknowns which gives us three sections, because the sum  $x$  is known from observations of direct sunlight. In the following we have also introduced the ozone in a bottom section  $u$  as being known from direct measurements, and have thus four sections within each of which the amount of ozone is assumed to be uniform at different heights or five sections, if we include the highest section where the amount of ozone is assumed negligible.

Let the ozone in the different sections be represented as follows (see fig. 4) :—

Above 50 km. = 0 cm.

Between 50-35 km. =  $x_1$

Between 35-20 km. =  $x_2$

Between 5- 2 km. =  $u$

Between 20- 5 km. =  $x - (x_1 + x_2 + u)$ .

For example, in the middle section, a beam of 1 km. length has to traverse  $x_2/15$  cm. of ozone, while in the bottom section it would traverse  $u/3$  cm. of ozone. In the numerical integration we proceed in steps of 1 km. and make allowance for the light scattered above 50 km.

Now let

$L$  be the whole mass of air which the beam has to traverse;  $L$  is given by the formula below,

$l$  its path through the ozone measured in centimetres,

$\beta$  the scattering coefficient of the whole atmosphere (for 760 mm. of Hg).

$\alpha$  the absorption coefficient of ozone,

$b$  the pressure at the height of scattering (in millimetres of Hg).

Then we have

$$I = \Sigma 10^{-\beta L} . b . 10^{-\alpha l} \times \text{const.}$$

the summation in practice being carried out in kilometre steps from 2 km. to 65 km. It may be easily shown that

$$L = 1 + b/760 (f(Z) - 1),$$

where  $f(Z)$  tends to  $\sec Z$  as  $Z$  tends to 0 and is given, for example, in the tables of Bemporad.

The path  $l$  is a function which contains the unknowns  $x_1$  and  $x_2$ , and also  $x$  and  $u$ , and is found trigonometrically. The computation is made in steps

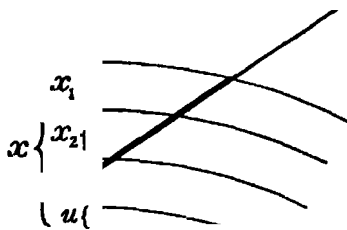


FIG. 4.

of 1 km., and the results are collected as the average of each layer. As an example the mean paths for  $Z = 80^\circ$  and  $Z = 90^\circ$  are given in Table IV.

Table IV.

Height.	Mean path $l$ in cm. of ozone.	
	For $Z = 80^\circ$ .	For $Z = 90^\circ$ .
Above 50 km.	$x$	$x$
50-35 km.	$x + 2.31x_1$	$x + 19.03x_1$
35-20 km.	$x + 4.38x_1 + 2.31x_2$	$x + 15.30x_1 + 18.90x_2$
20-5 km.	$3.31x + 1.75x_1 + 2.06x_2 - 2.31u$	$19.87x - 9.99x_1 - 3.62x_2 - 18.87u$
5-2 km.	$5.51x - 0.62x_1 - 0.33x_2 - 2.14u$	$22.34x - 13.18x_1 - 10.64x_2 + 17.52u$

Now we can form five expressions for  $10^{-u}$  and similarly five expressions for  $10^{-\beta L}$ . *b.* We have assumed the ozone per kilometre of the bottom section to be 1% of  $x$ .

When comparing the observed and calculated results, it should be noted that the instrument measures  $(-\log I/I' + C)$  where  $C$  is an unknown constant. For  $Z = 0^\circ$ ,  $(-\log I/I' + \text{another constant})$  can be computed if we know only  $x$  without knowing the distribution of the ozone. The unknown constants are eliminated by making the (extrapolated) observed value at  $Z = 0^\circ$  equal to the calculated value here. For instance, for  $Z = 0^\circ$  when  $x = 0.280$  cm. the observed (extrapolated) value of  $-\log I/I'$  is 0.338, while the computed value is 0.425; the calculated values are therefore reduced by 0.087. (For  $x = 0.200$  cm. the reduction is 0.117, see "Secondary Scattering," § VII.)

If  $x$  and  $u$  are known, we have to evaluate two unknowns  $x_1$  and  $x_2$ . This requires two equations. For this we make use of the results for  $I/I'$  for  $Z = 90^\circ$  and  $Z = 80^\circ$ . Of course it will be desirable also to choose a third value of  $Z$  in order to have a check; then we choose  $Z = 86.5^\circ$ .

As an instance we give the equation for  $x = 0.280$  cm. and  $Z = 80^\circ$ :

$$\begin{aligned}
 &68.50 \cdot 10^{-0.08(-0.62x_1 - 0.33x_2 + 1.52)} + 445.9 \cdot 10^{-0.08(1.75x_1 + 2.06x_2 + 0.92)} \\
 &\quad + 107.58 \cdot 10^{-0.08(4.38x_1 + 2.31x_2 + 0.28)} + 12.19 \cdot 10^{-0.08(2.31x_1 + 0.28)} \\
 &\quad + 1.35 \cdot 10^{-0.08 \cdot 0.28} \\
 &- 22.39 [30.4 \cdot 10^{-1.275(-0.62x_1 - 0.33x_2 + 1.52)} + 291.47 \cdot 10^{-1.275(1.75x_1 + 2.06x_2 + 0.92)} \\
 &\quad + 84.89 \cdot 10^{-1.275(4.38x_1 + 2.31x_2 + 0.28)} + 9.87 \cdot 10^{-1.275(2.31x_1 + 0.28)} \\
 &\quad + 1.103 \cdot 10^{-1.275 \cdot 0.28}] = 0.
 \end{aligned}$$



This equation corresponds to a curve in which  $x_1$  is plotted against  $x_2$ . A further equation ( $Z = 90^\circ$ ) gives a second curve and the intersection is the numerical solution of the two simultaneous equations.

Fig. 5 gives the curves for  $x = 0.280$ . It is notable that for low sun  $x_1$  is nearly independent of  $x_2$  as the curve is approximately a straight line parallel to the axis of  $x_2$ . That means that at lowest sun the reading of the instrument is determined only by the ozone in the highest layers.

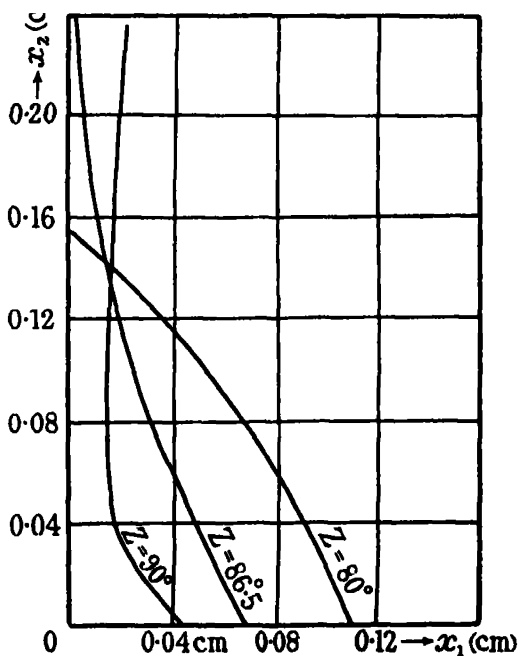


FIG. 5.

(b) *Results of Method A.*—Table V gives the results for three different values of the total ozone content in centimetres per section.

Table V.—Cm. of Ozone in Different Sections of the Atmosphere.

Range.	$x = 0.200$ cm.	$x = 0.280$ cm.	$x = 0.340$ cm.
	cm.	cm.	cm.
50–35 km.	0.016	0.016	0.017
35–20 km. . . . .	0.123	0.144	0.172
20– 5 km. . . . .	0.055	0.112	0.141
5– 2 km.	(0.006)	(0.008)	(0.010)

The height of the centre of gravity is 23.9 km. for  $x = 0.200$ ; 21.5 km. for  $x = 0.280$ ; and 21.3 km. for  $x = 0.340$ . The centre of gravity is higher, when the total amount of ozone is low.

The "block" distributions are shown in fig. 6 and smooth curves have been drawn for each distribution, keeping the amount of ozone in each section equal to that in the corresponding "block." Table VI also gives the distribution in centimetres of ozone per kilometre of height as obtained from these smooth curves, and also the ratio of ozone to air.

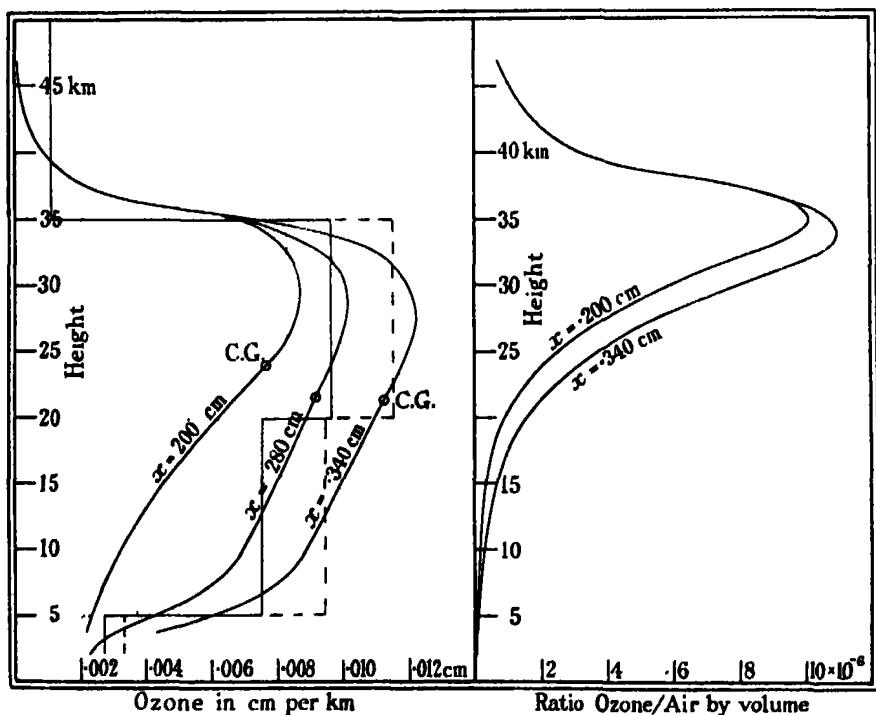


FIG. 6.—Distribution of ozone, (a) in cm. per km. of height, and (b) as ratio of ozone to air. (Method A.)

#### V. Method B of Calculating the Vertical Distribution of Ozone.

(a) The second method of finding the ozone distribution consists in dividing the atmosphere into a number of horizontal sections such that the mass of the air in any section is  $1/\sqrt{10}$  that of the section below it. The whole of the air and ozone in any section is then supposed to be concentrated in a very thin layer at a suitable height above the bottom of the section. Let us consider the light scattered from the direct solar beam into the instrument by any

layer; this light will have suffered absorption by ozone and reduction due to scattering by air in passing obliquely through the layers above the scattering layer, and through half the air and ozone in the scattering layer. After scattering, the light suffers further absorption by ozone and reduction due to scattering by the air on the remaining vertical part of its path to the instrument. The total light received by the instrument is the sum of that scattered by all the layers. The amount of scattering is assumed to be proportional to the mass of air in the section and inversely as the fourth power of the wave-length.

Making any assumption with regard to the ozone distribution it is possible to calculate the scattering and absorption by the different layers for any value

Table VI.—Distribution of Ozone by Method A. Cm. of ozone per km. of height; and ratio by volume of ozone to air.

Height above sea-level.	$x = 0.200$ cm.		$x = 0.280$ cm.		$x = 0.340$ cm.	
	Cm. $\times 10^{-3}$ .	Ratio $\times 10^{-6}$ .	Cm. $\times 10^{-3}$ .	Ratio $\times 10^{-6}$ .	Cm. $\times 10^{-3}$ .	Ratio $\times 10^{-6}$ .
km.						
45	0.15	1.00	0.15	1.00	0.15	1.00
40	0.87	3.0	0.87	3.0	0.87	3.0
35	6.95	10.2	7.08	10.4	7.28	10.7
30	8.7	5.8	10.1	6.8	12.0	8.0
25	8.2	2.5	9.9	3.1	12.0	3.7
20	6.1	0.87	8.9	1.3	11.0	1.6
15	4.4	0.30	8.0	0.53	10.0	0.66
10	3.2	0.10	7.0	0.22	8.9	0.28
5	2.3	0.04	4.2	0.08	6.0	0.11
2	—	—	2.2	0.03	—	—

of the zenith distance for each of the two wave-lengths and so to calculate an "Umkehr Curve" which may be compared with the observed curve, except for certain constants.

The procedure is then to calculate "Umkehr Curves" on the basis of some assumed ozone distribution until a curve agreeing with the observed curve is obtained. Actually a certain standard ozone distribution having certain percentages of the total ozone in each layer is first taken, and "Umkehr Curves" are calculated for different values of the total ozone content. Afterwards the effects are found of small modifications of this distribution in which a definite percentage of the total ozone is transferred from one layer to a neighbouring one. Once these laborious calculations are complete, it is easy to calculate an "Umkehr Curve" for any assumed ozone distribution—not greatly different from the standard distribution—and for any value of the total ozone.

(b) *Assumed Vertical Temperature and Pressure Distribution in the Atmosphere and the Heights of Sections and Layers.*—In order to determine the thickness of the sections into which the atmosphere is divided for the purpose of calculation, it is necessary to assume a vertical temperature distribution. Up to 20 km. the average value is known accurately, and above this height the temperature has been taken from Whipple's results from air waves from gunfire. Any error introduced here will have only a small effect. The actual values assumed are given in Table VII.

Table VII.

Height above sea-level (km.)	0	2	10	12	20	30	32	40	50
Temperature in ° Abs.	280	270	230	220	220	220	220	275	336
Pressure in mm.	760	590	198	147	42.7	9.12	6.74	2.19	0.78
Dens./dens. at N.T.P.	0.975	0.782	0.312	0.240	0.0697	0.0140	0.0110	0.00285	0.00086

The heights of the boundaries between the sections, and the heights of the layers into which the air and ozone is supposed to be concentrated are given in Table VIII.\*

Table VIII.

Number of layer or section	1	2	3	4	5	6	7	8
Height of bottom of section above M.S.L. (km.)	1.9	10.2	17.8	25.2	32.7	41.2	52.8	65.3
Height of layer above M.S.L. (km.)	4.9	13.0	20.5	27.9	35.4	45.3	57.0	69.9

The position of the hypothetical layer within any section of the atmosphere is taken at such a height that half the air in the section is above this level and half below it. This is correct so long as the extinction of light within the section is small. When the sun is very low the theory must be modified (see § V (e)).

V (e). *Details of Method B.*—Let fig. 7 represent a vertical section through the atmosphere and let the lines 1, 2, 3, ..., represent the thin layers into which the atmosphere is supposed to be concentrated. Let the air in layer 1 have an extinction coefficient due to scattering by air molecules and fine dust,  $\beta_1$  for the wave-length 3110 Å., and  $\beta'_1$  for the wave-length 3290 Å. and similarly for layer 2, etc. Let  $x_1$  be the amount of ozone in layer 1 (measured, as usual, in centimetres thickness of pure ozone at N.T.P.) and similarly for layer 2, etc. Let  $\alpha$  be the absorption coefficient of ozone for the wave-length 3110 Å. and  $\alpha'$  for the wave-length 3290 Å.

\* A correction is made in the calculations for the presence of air above section 8, and it has been shown that the amount of ozone above section 8 is negligible.

Consider any ray OA from the sun which is supposed to be scattered by layer 3, part of the scattered light passing vertically down to the instrument. Let Z be the angle at which this ray meets layer 3, and let  $\zeta_{34}$ ,  $\zeta_{35}$ ,  $\zeta_{36}$ , etc., be the angles at which it meets layers 4, 5, 6, etc.

Let S be the flux of radiation of wave-length 3110 Å. falling on the outside of the atmosphere in the direction OA. Then the flux reaching the centre of layer 3 will be given by

$$\log F_3 = \log S - (\alpha_3 + \beta_3) \sec \zeta_{33} \dots (\alpha_4 + \beta_4) \sec \zeta_{34} - \frac{1}{2} (\alpha_3 + \beta_3) \sec Z. \quad (1)$$

Of this flux some fraction  $f_3$  (where  $f$  is proportional to the density of the air) will be scattered towards the instrument so that after absorption on the vertical path from layer 3 to the ground, the flux  $I_3$  received by the instrument will be given by

$$\log I_3 = \log F_3 + \log f_3 - \frac{1}{2} (\alpha_3 + \beta_3) - (\alpha_2 + \beta_2) - (\alpha_1 + \beta_1). \quad (2)$$

The total flux received by the instrument from the clear blue zenith sky for the wave-length 3110 Å. will then be

$$I = I_1 + I_2 + I_3 \dots + I_6,$$

and similar equations will hold for the wave-length 3290 Å. Thus we may determine the value of  $I/I'$ .

V (d). *Determination of Constants.*—

The readings of the instrument give the value of  $-\log I/I'$  plus an unknown constant. This constant arises from such effects as the spectral sensitivity of the photoelectric cell: slit widths, etc. These could be measured but we still should not know  $S/S'$ , the ratio of the energies in the two wave-lengths as received at the outside of the atmosphere, so that there is no object in measuring them. We must, therefore, find some method whereby the calculated and observed values may be linked together and made comparable. For this purpose use is made of the fact that

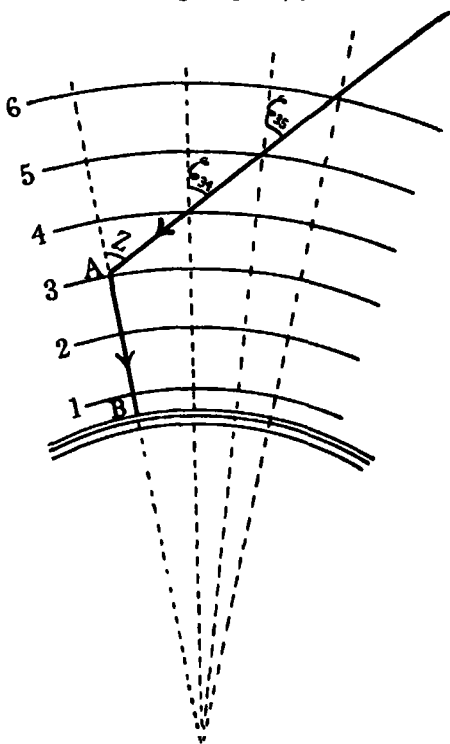


FIG. 7.

when the zenith distance is less than  $40^\circ$  the value of  $-\log I/I'$  does not depend appreciably on the distribution of the ozone but only on the total quantity. We therefore compare the calculated values of  $-\log I/I'$  when  $Z = 37.5^\circ$  with the observed value\* and add the difference to the calculated values of  $-\log I/I'$  for all values of  $Z$ . Thus the observed and calculated values are made to agree when  $Z = 37.5^\circ$ . The dotted curve of fig. 2 ( $Z = 37.5^\circ$ ) represents the calculated values of  $-\log I/I'$ . It will be seen that this curve does not differ by a constant from the observed curve. The reason for this will be discussed in § VII (c) (Secondary Scattering).

V (e). *Sunset Theory*.—The general treatment of method B does not allow one to calculate  $-\log I/I'$  accurately when  $Z$  becomes nearly  $90^\circ$ , i.e., at about sunset or sunrise. The reason for this is that it is not sufficiently accurate to say that the absorption by ozone and air in the layer ( $r$ ) which is being considered as scattering light into the instrument is

$$(\alpha x_r + \beta_r) (1 + \sec Z)/2,$$

since this expression becomes infinite when  $Z = 90^\circ$ . In addition, it is not sufficiently accurate to assume that the scattering layer is placed within the section which it represents in such a way that half the air is above it and half below; hence values of  $\zeta_{rs}$  must be calculated by a better method. When  $Z$  is greater than  $87.1^\circ$ ,  $-\log I/I'$  calculated by the general treatment has an error of more than 0.002, and this error rapidly increases as  $Z$  is further increased. As calculations for  $Z = 90^\circ$  yield valuable information it was necessary to make a better approximation. The integral (1) of § IV was therefore evaluated graphically over the range in which method B breaks down. In practice this range was found to correspond to whichever section of the atmosphere was considered as scattering the light. Thus it was necessary to replace the terms  $-(\alpha x_3 + \beta_3) \sec Z/2 - (\alpha x_3 + \beta_3)/2$  in equations (1) and (2) of § V (c) by  $-(\alpha x_3 + \beta_3) \phi(Z)$ ; so that the combined equations became:

$$\begin{aligned} \log I_3 = \log S - (\alpha x_3 + \beta_3) \sec \zeta_{33} - \dots - (\alpha x_4 + \beta_4) \sec \zeta_{34} \\ - (\alpha x_3 + \beta_3) \phi(Z) - (\alpha x_2 + \beta_2) - (\alpha x_1 + \beta_1), \end{aligned}$$

for scattering in layer 3, and similarly for scattering in any other layer; where  $\phi(Z)$  is a function which can be evaluated by graphical integration. The values of  $\zeta_{rs}$  were also calculated.

\* The observed value of  $-\log I/I'$  is obtained from the smooth curve of fig. 2 ( $Z = 37.5^\circ$ ).

Values of the functions  $\phi(Z)$  and  $\zeta_{rs}$  were found for  $Z = 90^\circ$  and  $Z = 91.1^\circ$ . On applying this technique (called the "sunset theory") to observations taken at and after sunset it was found that the values of  $-\log I/I'$  were very insensitive to the ozone in layers 5 and under. Hence the sunset theory determines the values of  $x_6$  and  $x_7$ , almost independently of the values of  $x_8$ ,  $x_4$ , etc. The points  $Z = 90^\circ$  and  $91.1^\circ$  are not sufficiently far apart to determine both  $x_6$  and  $x_7$ , and  $91.1^\circ$  is the greatest value of  $Z$  at which a group of consistent observations could be obtained. Thus while the observational data are sufficient to determine the amount of ozone above layer 5, the distribution of ozone above layer 5 cannot at present be determined.

The sunset theory, then, determines the ozone above layer 5, provided that  $x_5$ ,  $x_4$ , etc., are known approximately from the general treatment of method B, and provided that  $x$  is known from direct sun observation. Similarly the general treatment determines  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$ , and  $x_5$ , if a small correction is made after application of the sunset theory, for the exact value of  $x_6$ . Since  $x$  is known independently the general treatment only evaluates four new unknowns. Consequently the six points, apart from  $Z = 37.5^\circ$ , at which agreement is sought with the observed "Umkehr Curve" are more than sufficient for the purpose.

## VI. Results of Method B.

(1) *Distribution of Ozone with Height.*—The chief difficulty in applying method B successfully arises from the fact that very great accuracy of observation is required, if the distribution of ozone is to be calculated with any reasonable accuracy. The evaluation, by method B, of a distribution from a given "Umkehr Curve" consists effectively of obtaining the most probable solution of six simultaneous equations involving four unknowns. The six simultaneous equations arise from the six\* points (at  $Z = 70.5^\circ$ ,  $80.4^\circ$ ,  $83.65^\circ$ ,  $84.85^\circ$ ,  $86.07^\circ$ ,  $87.11^\circ$ ) on the observed "Umkehr Curve" at which values of  $-\log I/I'$  are also calculated by method B; the four unknowns are associated with the ozone contents of layers 1, 2, 3, 4, and 5, the sum being known.

The most probable solution of six simultaneous with four unknowns is the "least square solution" such that the sum of the squares of the residuals is a minimum. The rigid determination of a large number of least square solutions would take too long, and therefore a method of trial and error, which is much shorter, was normally used. Unfortunately the probable errors of

\* There would be very little improvement in the accuracy of results if more than six points were taken, because more than six points would be much more interdependent.

values of  $x$ , obtained from a single "Umkehr Curve" are so large that the resulting ozone distribution is almost meaningless. The even ozone "Umkehr Curves," an account of which has been given in § I, were therefore used as observational data for the evaluation of distributions. While these calculations of the distribution of ozone give accurate results for the average height, the standard error of the distribution which they give is so large that they are still of little value for finding the real distribution. Still greater accuracy of observational data was obtained by making the assumption that the vertical distribution of ozone changes in a uniform, regular manner, as the total ozone content changes. As a result, the change in distribution, assumed uniform, could be found by obtaining theoretical " $-\log I/I', x$ " curves whose slope agreed with the observed points given in fig. 2. The theoretical curves corresponding to the final distribution shown in fig. 8 and Table X are also given by the lines in fig. 2. The agreement may be considered good enough to justify the assumption of uniform change in distribution, but does not constitute a proof that the assumption was correct.

The usefulness of this method of treatment is that it enables an observed "Umkehr Curve" at any value of  $x$  to be adjusted to what it would be at some given value of  $x$ , say  $x = 0.260$  cm. Consequently it was possible to find the mean of the 46 "Umkehr Curves" available, after adjusting them to the

Table IX.—Distribution of Ozone (Method B) "Least Square" Solution of "General Mean Umkehr Curve." (Cm. of ozone per layer.)

Layer.	Height above sea-level.	Total ozone content (cm.).						Standard error.
		0.200.	0.220.	0.240.	0.260.	0.280.	0.300.	
	km.							cm.
6	45.1	0.0039	0.0040	0.0042	0.0042	0.0043	0.0045	0.0008
5	35.9	0.0306	0.0319	0.0329	0.0338	0.0343	0.0344	0.0016
4	28.1	0.079	0.082	0.085	0.087	0.088	0.089	0.014
3	20.6	0.025	0.031	0.037	0.044	0.051	0.058	0.032
2	13.1	0.034	0.042	0.049	0.057	0.067	0.077	0.034
1	4.9	0.028	0.030	0.032	0.034	0.035	0.037	0.014

value  $x = 0.260$  cm. This "General Mean Umkehr Curve" was taken as the basis of further calculations by method B, and a rigid least square solution was found. The results of this solution are given in the column  $x = 0.260$  of Table IX. The standard errors of the solution are given in the last column of the same table. Table IX also gives the ozone contents of different layers



for various values of the total ozone, on the assumption, stated above, that the distribution changes uniformly with the total ozone.

Table IX gives the amounts of ozone which would be required in six thin layers, at the fixed heights given in Table VIII, in order to give the same "Umkehr Curve" as is actually observed. The true distribution of the ozone in the atmosphere must naturally be continuous, and it is necessary to consider how a smooth curve, giving the amount of ozone per kilometre of height, should be obtained from the six thin layers of ozone.

In fixing the positions of the thin hypothetical layers in the atmosphere it has been found convenient to divide the atmosphere so that the pressure at the bottom of each section is  $\sqrt{10}$  times the pressure at the top; and the layers are taken so that half the air of a section is above and half below the height of the layer. For convenience the ozone has been taken as concentrated into thin layers at the same height as the layers of air, but it would not have affected the smooth distribution of ozone, which we wish to obtain, if quite different positions had been taken for these hypothetical layers. Thus, in order to obtain the amount of ozone per kilometre of height it is not correct to divide the ozone in each layer by the thickness of the section of air corresponding to that layer; it is necessary to draw the smooth curve very carefully taking care that each layer of ozone is replaced by a section of ozone which has the same centre. If this is not done it will be found that the height of the centre of gravity of all the ozone given by the smooth curve is not equal to the height given by the discrete layers.

The amount of ozone per kilometre of height at different levels is given in Table X for different values of the total ozone content, and is shown graphically in fig. 8. Fig. 8 and Table X also give the ratio of ozone to air at different heights. It will be seen that the unevenness in layers 2 and 3 in Table IX does not correspond to an uneven distribution of ozone in the actual atmosphere.

(2) *Mean Height of the Ozone.*—The heights of the individual layers are known, and by taking moments of the values of  $x$ , about the surface of the earth the mean height of the ozone can be found. The standard error of a determination of  $h$  from a single day's observations is found to be about  $\pm 1.5$  km. This is so small that the elaborate methods of obtaining the means of the observations used above are not necessary in the determination of  $h$ . To avoid the necessity of a large number of individual determinations of  $h$ , however, the "even ozone Umkehr Curves," Table I, have been used, rather than separate day's results. The effect of this has been to reduce slightly the accuracy of the determined values of  $h$ , because as was shown in § II the shape

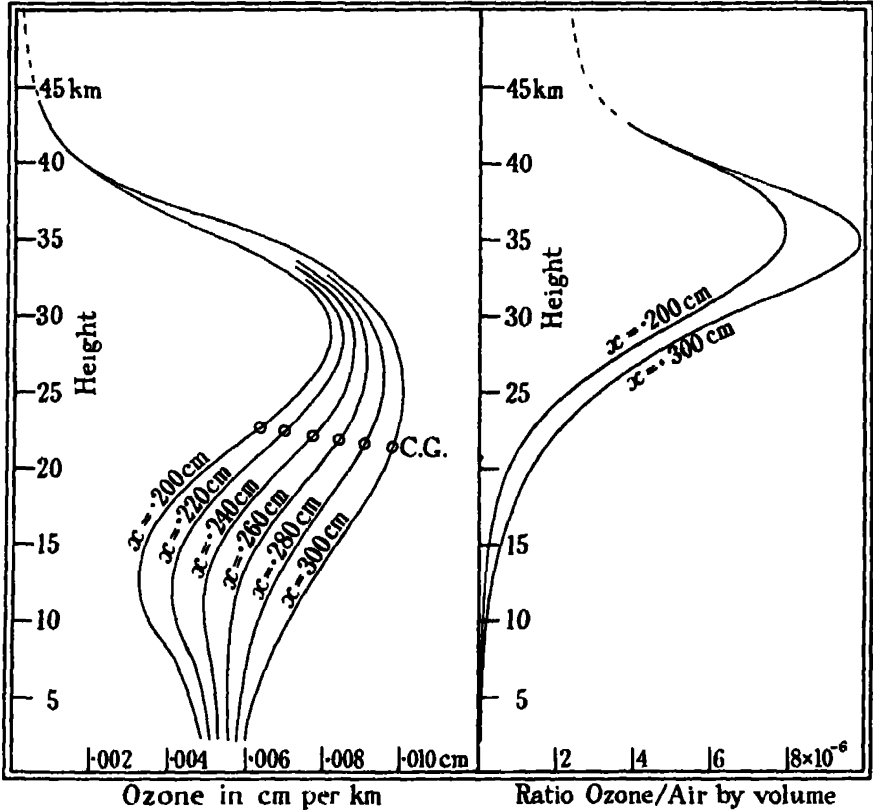


FIG. 8.—Distribution of ozone, (a) in cm. per km. of height, and (b) as ratio of ozone to air. (Method B.)

Table X.—Distribution of Ozone by Method B. (Cm. of ozone per km. of height ; and ratio by volume of ozone to air.)

Height above M.S.L.	$x = 0.200.$		$x = 0.220.$		$x = 0.240.$		$x = 0.260.$		$x = 0.280.$		$x = 0.300.$	
	Cm. $\times 10^{-3}$	Ratio $\times 10^{-4}$	Cm. $\times 10^{-3}$	Ratio $\times 10^{-4}$	Cm. $\times 10^{-3}$	Ratio $\times 10^{-4}$	Cm. $\times 10^{-3}$	Ratio $\times 10^{-4}$	Cm. $\times 10^{-3}$	Ratio $\times 10^{-4}$	Cm. $\times 10^{-3}$	Ratio $\times 10^{-4}$
km.												
50*	0.20	2.3	0.20	2.3	0.20	2.3	0.20	2.3	0.20	2.3	0.20	2.3
45	0.42	2.8	0.42	2.8	0.42	2.8	0.42	2.8	0.42	2.8	0.42	2.8
40	1.73	6.1	1.73	6.1	1.73	6.1	1.73	6.1	1.73	6.1	1.73	6.1
35	5.40	7.9	5.60	8.3	5.93	8.7	6.19	9.1	6.46	9.5	6.72	9.9
30	8.1	5.4	8.4	5.6	8.6	5.8	8.9	6.0	9.1	6.1	9.4	6.3
25	7.4	2.2	8.0	2.4	8.5	2.6	9.0	2.7	9.5	2.9	10.1	3.0
20	5.0	0.72	6.0	0.85	6.9	0.99	7.8	1.1	8.7	1.2	9.6	1.4
15	3.4	0.23	4.4	0.29	5.3	0.35	6.3	0.42	7.3	0.48	8.3	0.55
10	3.5	0.11	4.2	0.13	5.0	0.16	5.7	0.18	6.4	0.21	7.1	0.23
5	4.6	0.08	4.9	0.08	5.3	0.09	5.6	0.10	5.9	0.11	6.2	0.11
2	4.9	0.06	5.1	0.07	5.4	0.07	5.6	0.07	5.8	0.07	6.0	0.08

\* Ozone values at this height are approximate.

of the observed "Umkehr Curve" is not strictly dependent on  $x$  only. The results are given in Table XI below.

The values of  $h$ , using the "general mean Umkehr Curve" agree very well with the above, thus providing some confirmation of the validity of the assumption that the distribution of the ozone varies uniformly with  $x$ .

The above values of the height of the ozone have a high order of accuracy. The probable error of  $h$  should be much less than the probable errors of  $x$ ; for if in calculating the values of  $x$ , from a given set of observations we try to increase, e.g.,  $x_2$  at the expense of  $x_1$ , then to obtain again a good fit with the observed "Umkehr Curve," we find it necessary to increase  $x_2$  further, this

Table XI.—Average height of ozone above M.S.L. (km.).

Data used.	Method.	$x =$ 0.200 cm.	$x =$ 0.220 cm.	$x =$ 0.240 cm.	$x =$ 0.260 cm.	$x =$ 0.280 cm.	$x =$ 0.300 cm.	$x =$ 0.340 cm.
Even ozone Umkehr Curves "	A	23.9	—	—	—	21.5	—	21.3
	B	22.3	22.6	22.2	21.9	21.4	21.9	—
General mean Um- kehr Curve "	B	22.6	22.4	22.1	21.8	21.5	21.3	—

Standard error of height = 0.2 km.

time at the expense of  $x_2$ . Hence the effect on  $h$  of quite large changes in  $x_1$ ,  $x_2$ , and  $x_3$  is found to be very small, provided that a reasonably good agreement with the observations is maintained.

## VII. Estimate of Accuracy of Results.

(a)  $x$  from Direct Sun Observations.—The question of the standard error in  $x$  as obtained from direct sun observations was discussed in § II. The standard error of the mean ozone content for a morning or afternoon was found to be  $\pm 0.002$  cm.; but the standard errors of observations given in Table III would only result in a standard error in ozone at the time of the observations of  $\pm 0.0008$  cm. In addition to the error of observation there is the possibility of errors in the constants of the apparatus, which would produce persistent errors in  $x$ .  $\log I_0/I_0'$ , discussed in § III, was found to have a standard error of  $\pm 0.003$ ; an error in  $\log I_0/I_0'$  of 0.003 would produce a persistent error in  $x$  of  $0.003/\sec \zeta_h$  cm. Errors in  $\alpha$  and  $\alpha'$  would produce a persistent error in  $x$  such that the error in  $(\alpha - \alpha')x$  is zero. Therefore wherever  $(\alpha - \alpha')x$  occurs, errors in  $\alpha$  and  $\alpha'$  do not matter, and in calcu-

lations of height and distribution of ozone  $(\alpha - \alpha')x$  are always associated with one another. The other quantity which is subject to appreciable error is  $(\beta - \beta')$ ; the Rayleigh value of  $(\beta - \beta')$  for pure air for the altitude of Arosa is 0.060; the value 0.065 has been adopted, as being reasonable for the air and dust at Arosa. (In method A  $\beta - \beta' = 0.090$  for sea-level has been used.) An error of  $\pm 8\%$  in  $(\beta - \beta')$  would cause an error of  $\pm 0.006$  cm. in  $x$ ; we believe that the actual error in  $(\beta - \beta')$  is always less than this.

In order to obtain a value of  $x$  from direct sun measurements it is necessary to assume that  $x$  has no regular diurnal variation, symmetrical about noon. If this assumption were wrong there would be no very serious error; it can be shown that if there is such a diurnal variation of  $x$ , the value of  $x$  obtained at any time of day from direct sun measurements, on the assumption of no variation in  $x$ , will be the true value of  $x$  near sunrise or sunset.

(b) *Height from Direct Sun Observations* is subject to a standard error of approximately  $\pm 5$  km. This is so large that the error in  $h$  due to a faulty value of  $(\beta - \beta')$  is negligible. If there were really a diurnal variation of  $x$ , its sunset value being different from the noon value by  $\pm 10\%$ , then  $h$  by direct sun methods would be wrong by  $\pm 12$  km.;  $h$  would be 12 km. too high if a diurnal variation of  $x$  had been neglected in which  $x$  was lower at noon than at sunrise or sunset.

(c) *Errors in Distribution and Height, from Zenith Blue Sky Observations (Method B).*—The standard error of the distribution was given in Table IX, while that of the height is 0.2 km. These errors arise from the following causes :—

Errors in  $x$  as determined from direct sun observations.

Errors due to variation of  $x$  or  $(\beta - \beta')$  during the day.

Errors due to the assumption that  $(\beta - \beta')$  is always the same from day to day.

Random errors of observation of  $-\log I/I'$  for zenith blue sky observations.

Persistent errors have not been included. If  $(\beta - \beta')$  has been persistently given a value which is 8% too low, then the value of  $h$  found will be from 0 to 0.6 km. too low, depending on how the extra haze is distributed with height. The maximum error in  $h$  occurs if all the extra haze is at the surface; but if the extra haze is distributed like the ozone it will produce no error in  $h$ , for it will have the same effect as a persistent error in  $(\alpha - \alpha')$ , and this is known to have no effect on  $h$ . If there is a real symmetrical variation of  $x$  throughout the day, such that  $x$  is 10% lower at noon than at sunrise or sunset, then

the value of  $h$  as obtained above will be 0.75 km. too low. A real variation of 5% in  $(\beta - \beta')$  throughout the day, which has a similar effect to that of a real variation of  $x$ , would only produce an error in  $h$  of  $\pm 0.1$  km.

The above sources of persistent error produce an error in distribution which is small compared with the random errors.

*Secondary Scattering.*—The one remaining important source of persistent error is the fact that in the theory, light which is scattered into the instrument from direct sunlight only has been considered, and not light which is scattered into the instrument from the diffuse radiation within the atmosphere. No attempt has been made to deal with light which is scattered more than once on its way into the instrument. If the density of the ozone can be assumed constant in the region where the diffuse radiation is most intense, it can be shown that for each of the wave-lengths 3110 and 3290 Å. the ratio of light entering the instrument which has been scattered more than once to light which has been scattered once only is very nearly a constant independent of  $Z$ . For light of 3290 Å. the secondary scattered light is very nearly one-half of the total; for light of 3110 Å. less than one-half.

At first sight it seems that by neglecting secondary scattering a large error has been introduced into the calculations; but it was shown in § IV that any factor in the integral (1) which either does not affect  $-\log I/I'$ , or adds to it a constant independent of  $Z$  and  $x$ , may be neglected. Although secondary scattering affects  $I$  and  $I'$  very greatly, its effect on  $-\log I/I'$  is to change it by an amount which is very nearly independent of  $x$ . This is evident from fig. 2 ( $Z = 37.5^\circ$ ) which is a plot of  $-\log I/I'$  against  $x$  for the constant value of zenith distance  $Z = 37.5^\circ$ . It can be shown that the value of  $-\log I/I'$  when  $Z = 37.5^\circ$  is independent of the distribution of ozone, so that the difference in slope between the dotted line, which is calculated neglecting secondary scattering, and the discrete points which are plots of observations, is a measure of the importance of secondary scattering. We should expect the inclusion of secondary scattering to diminish the calculated value of  $I$  more than  $I'$ , so that  $-\log I/I'$  would be increased at all values of  $x$ . At  $x = 0.300$  cm. the increase would be greater than at  $x = 0.200$  cm.; hence the slope of the calculated curve would be greater. Thus the neglect of secondary scattering is of the right sign to account for the discrepancy between the slope of the calculated line and the observed points. (N.B.—Owing to the neglect of various constants the calculated values of  $-\log I/I'$  are known only, plus an unknown constant which is independent of  $x$  and  $Z$ ; hence we can only compare the slope of the calculated line with the observed points, and not the

position.) Allowance has been made in the calculations for the change of the constant with  $x$ .

An approximate investigation has been made into the magnitude of the correction to the calculated values of  $-\log I/I'$  for  $Z = 37.5^\circ$  which is necessitated by secondary scattering. The mean free path in pure ozoneless air at N.T.P. of a photon of wave-length 3110 Å. is about 12 km. The presence of ozone which strongly absorbs light of 3110 Å. would be expected to reduce this mean free path by absorbing more of the photons which have a long free path. The reduction would be the greater the more the ozone in the region of greatest density of the diffuse radiation, i.e., in the region near the surface of the earth. We believe there is rather more ozone near the earth's surface when  $x = 0.300$  cm. than when  $x = 0.200$  cm. Calculated values of  $-\log I/I'$  would agree very well with the observed points for  $Z = 37.5^\circ$  if the mean free path of a photon of 3110 Å. were 9.5 km. when  $x = 0.200$  cm., and 9.0 km. when  $x = 0.300$  cm. These figures are reasonable, and show that the correction to  $-\log I/I'$  for secondary scattering is of the right magnitude for agreement with observation, as well as of the right sign.

It is probable that even in the presence of ozone the secondary scattered light entering the instrument bears a nearly constant ratio to the primary scattered light, and this ratio is nearly independent of  $Z$ , for each wave-length. A test has been made of the effect on the calculated distribution and height, if the effect of secondary scattering were twice as great for low sun as for high sun. A table, XII, of the results of this assumption is given compared with the results of neglecting secondary scattering.

Table XII.

Layer.	1.	2.	3.	4.	5.	6.	Height.
% of total ozone per layer (method B)	15.5	17.5	19.5	33.5	13	1	km. 21.6
Ditto, assuming secondary scattering as above	18.5	17	19	33.5	11	1	20.9

We believe that the assumption made in calculating the above table is quite extreme.

(d) *Errors in Height and Distribution from Method A.*—In method A three observations are, in effect, used to determine two unknowns, while in method B six observations are used to determine four unknowns. Owing to the larger

number of layers into which the atmosphere is divided the details of the distribution can be found rather better by method B, but against this must be set the fact that the six observations used in method B are somewhat less independent than the three observations used in method A. Considering everything the errors of the two methods are probably very similar in magnitude and the standard errors found for method B can probably be applied to method A also.

### VIII. *Conclusions.*

We believe that the results obtained above give the general character of the vertical distribution of the ozone in the atmosphere in Switzerland, but we would call attention to the large standard errors of the amounts of ozone in the middle layers. The smooth curves showing the most probable distribution tend to give an impression of greater accuracy than they really have. On the other hand, the height of the centre of gravity of the ozone can be relied upon to within a kilometre. There is, as yet, no evidence regarding the distribution of ozone in other parts of the world though such information would be of great interest.

One might expect that at the higher levels the ratio of ozone to oxygen would be constant and the apparent fall in this ratio, about 40 km., is rather surprising. Neither the uncertainty in the amount of ozone nor of oxygen at these levels would seem sufficient to account for this apparent fall as due to errors.

The results obtained by method B for the surface layers are much less accurate than the values obtained by direct measurement of the absorption of artificial light in traversing the air near the ground. They would, however, be reconcilable with them if the amount of ozone in the atmosphere decreased rapidly near the ground, as might be expected if it were destroyed by dust, etc.

The main changes in ozone content appear to be centred at a height of about 10 km. to 20 km., a fact of considerable meteorological interest.

The present figures for the distribution of ozone are so different from those previously assumed that it will be necessary to re-examine such questions as the temperature of the atmosphere at great heights. Though the amount of ozone in the upper layers is now thought to be smaller than was formerly assumed, it does not necessarily follow that the temperature will be lower than former estimates, for Götz has pointed out that the temperature above 40 km. may be governed more by the absorption of oxygen than by ozone.\*

\* Götz, 'Meteor. Z.', vol. 50, p. 455 (1933).

*Acknowledgments.*

We wish to thank the Royal Society for a grant covering the cost of the transport of the large instruments to Switzerland and back. We are also indebted to Sheffield Town Trust, Queen's College, Oxford, and the Department of Scientific and Industrial Research for grants to one of us. We have also to thank the Chur-Arosa- and Rhätische Bahn for special concessions.

*Summary.*

A method is described for finding the average height of the ozone in the atmosphere and the general character of its vertical distribution. This method uses spectroscopic observations of the light from the zenith blue sky while the sun is rising or setting. A year's observations have been made at Arosa and the height and distribution of the ozone have been found for the different seasons of the year, and under various meteorological conditions. It is shown that the average height of the ozone in Switzerland is about 22 km. above sea-level and that the ozone is distributed mainly between the ground and 35 km. The form of the vertical distribution appears to depend chiefly on the total amount of ozone present.

---



## *The Energies of Alpha, Beta, and Gamma Rays.*

By H. A. WILSON, F.R.S., Rice Institute, Houston, Texas.

(Received February 17, 1934.)

This paper contains a discussion of the energies of alpha, beta, and gamma rays on the theory of nuclear energies previously suggested by the writer.\*

It is assumed that the nucleus emits gamma rays with energies equal to multiples of  $3.85 \times 10^5$  electron volts and that these rays either lose energy to the electrons or gain energy from them before escaping from the atom. The gamma rays may also eject electrons as beta rays. The energy transfer may be supposed to take place between the electrons and the gamma rays, or between the electrons and the nucleus, at the moment of the emission of the gamma ray. There is no real difference between these two suppositions.

According to this theory there are no low energy gamma rays emitted by the nucleus so that the low energy beta and gamma rays, which are observed, must be emitted by the electronic system after it has been excited by previous gamma rays, of energy  $3.85n$ , from the nucleus. The unit  $10^5$  electron volts will be used throughout. Some low energy beta rays have been shown by Ellis and others to be produced in this way, as is well known. Ellis,† for example, has shown that there are a number of beta rays from radium B with energies nearly equal to  $K - 2L_1$ ,  $K - L_1 - L_2$ ,  $K - L_1 - M_2$ , and so on.

The possible energies of the electronic system of an atom of high atomic number are approximately equal to  $\sum_m N_m E_m$ ‡ where  $N_m$  is a positive integer equal to the number of electrons removed from the energy level  $E_m$  to a zero level or to outside the atom.

The removal of a K electron changes the effective atomic number for the other K electron by about 0.33 and so alters K by about 0.01, since K increases by about 0.03 when the atomic number is increased by unity. The effect on the L levels is about 0.005. In the same way the removal of an L electron only changes the L levels by about 0.002 so that it is necessary to remove about five to change L by 0.01. The effect on the M levels is very small since they are small. Thus the errors due to taking the possible energies equal to

\* 'Phys. Rev.,' November 15, 1933; 'Proc. Roy. Soc.,' A, vol. 144, p. 280 (1934).

† 'Proc. Roy. Soc.,' A, vol. 139, p. 336 (1933).

‡ It is scarcely necessary to point out that only certain values of the integers  $N_m$  are possible.

$\Sigma N_m E_m$  will usually be less than 0.01, which is not much greater than the probable errors of the beta ray energies.

If the electronic system has energy  $\Sigma N_m E_m$  we may suppose that a transition is possible to any other state with energy  $\Sigma N'_m E_m$  provided this is less than the initial energy. The energy emitted is then  $\Sigma (N_m - N'_m) E_m$  which should therefore be a possible energy value for a beta or gamma ray. Strictly speaking such a gamma ray is really an X-ray but here we shall continue to call it a gamma ray. If we put  $N_m - N'_m = N''_m$  the energy of the ray emitted is  $\Sigma N''_m E_m$  and in this expression some of the values of  $N''_m$  may be negative integers. Thus Ellis' lines like  $K - L_1 - L_2$  may be supposed due to a transition from an initial state with energy  $K$  to final state with energy  $L_1 + L_2$ , so that the difference is  $K - L_1 - L_2$ . When the electronic system is excited by such a high energy photon as one with energy 3.85 $\mu$  it seems reasonable to suppose that several electrons may be moved to higher levels.

It is unfortunately possible to represent almost any number by a summation like  $\Sigma N_m E_m$  so that almost any gamma or beta ray energy can be obtained by a suitable choice of the integers  $N_m$  and the levels  $E_m$ . It is therefore almost useless to show that the energies of the low energy beta rays can all be represented by  $\Sigma N_m E_m$ . There are, however, a number of cases where several adjacent energy values are nearly equal to very simple combinations of the energy levels and for these it seems probable that the agreements have real significance.

Table I gives nine adjacent low beta ray energies of radioactinium, taken from "Radiations from Radioactive Substances," by Rutherford, Chadwick, and Ellis.

Table I.

Beta ray energies.	$\Sigma N_m E_m$ .
0.834	$K - L_2 - M_1 = 0.837$
0.836	$K - 2M_1 = 0.839$
0.965	$K - 2M_2 = 0.961$
0.990	$K - M_1 = 0.991$
1.029	$K = 1.035$
1.074	$K + M_2 = 1.072$
1.106	$K + 2M_2 = 1.109$
1.135	$K + 2M_1 = 1.131$
1.305	$K + L_1 + 2M_2 = 1.301$

The last five values may be explained by supposing that the energy emitted is equal to the initial energy, so that the electronic system is left unexcited after the emission. The first three are values similar to those of Ellis mentioned

above. It seems rather difficult to believe that the above agreements are without significance.

Table II gives a rather similar set of beta ray energies from meso-thorium 2.

Table II.

Beta ray energies.	$\Sigma N_m E_m$ .
0.682	$K - 2L_1 = 0.684$
0.749	$K - 2L_2 + M_2 = 0.748$
0.791	$K - L_1 - 2M_2 = 0.792$
0.929	$K - L_2 = 0.930$
1.093	$K = 1.092$
1.129	$K + M_2 = 1.133$
1.245	$K + L_2 = 1.254$
1.279	$K + L_1 = 1.288$
1.337	$K + L_1 + M_2 = 1.337$
1.408	$K + L_1 + 2M_2 = 1.400$
1.644	$K + 2L_2 + L_1 = 1.646$

The second value, for example, can be explained by supposing that initially the atom has energy  $K + M_2$  and after the emission energy  $2L_2$ . This one only differs from one of Ellis' by having initial energy  $K + M_2$  instead of  $K$ . If  $L_1 + L_2$  is a possible final energy then there seems to be no reason why  $K + M_2$  should not be a possible initial energy. In the last two cases we have to suppose four electrons displaced initially. This may perhaps be considered unorthodox, but it should be remembered that a photon of energy  $3.85\mu$  emitted by the nucleus must produce an enormous intensity of radiation in the region occupied by the electrons. It is therefore not surprising that such multiple ionization or excitation should be produced, although nothing like it is usually observed in ordinary X-ray work. It is true that the representation of 1.644 by  $K + 2L_2 + L_1$  cannot be regarded as having much value as evidence, but this difficulty is inherent in the nature of the theory proposed.

Table III gives some of the low energy beta rays of thorium B. C.

Table III.

Beta ray energies.	$\Sigma N_m E_m$ .
0.983	$K + 2M_1 = 0.983$
0.989	$K + L_1 - L_2 + 2M_1 = 0.989$
1.107	$K + L_1 + M_1 = 1.106$
1.139	$K + L_1 + M_1 + M_2 = 1.138$
1.477	$2K - 2L_1 - M_1 = 1.493$
1.558	$2K - L_1 - L_2 = 1.563$
1.593	$2K - 2L_2 = 1.592$

The thorium line 0.989 above perhaps does not belong to thorium B. C, but to thorium C'' . Pb, since  $K + 3M_1 = 0.989$  for thorium C'' . Pb.

There are a number of low energy beta rays emitted by thorium B and its products which have not been assigned to any element. If we compare the energies of these beta rays with values of  $\Sigma N_m E_m$  for the different possible elements we can, in some cases, provisionally assign the lines to particular elements. For example, the beta ray energy 1.2699 is almost exactly equal to  $K + 2L_1 + M_1 = 1.2697$  for thorium B and the beta ray energy 1.2295 is nearly equal to  $K + 2L_1 = 1.2298$  for the same element. In the same way the beta ray energy 1.1856 is nearly equal to  $K + L_1 + L_2 = 1.1852$  for thorium C'' . Pb.

Small differences between the beta ray energies and  $\Sigma N_m E_m$  may be attributed to the omission of the N and O levels, and also to small errors in the values of the K, L, M levels besides possible small errors in the beta ray energies. Differences less than 0.005 may probably be ignored for such reasons.

Table IV.

Beta ray energies.		$\Sigma N_m E_m$ .
0.367	0.529 - $L_1 = 0.366$	$2L_1 + M_1 = 0.367$
0.374	0.529 - $L_2 = 0.372$	$L_1 + L_2 + M_1 + M_2 = 0.374$
0.396	0.529 - $L_3 = 0.395$	$L_1 + L_2 + M_1 + M_2 = 0.397$
0.444		$2L_1 + 2M_1 + M_2 = 0.443$
0.469		$L_1 + 2L_2 + M_2 = 0.468$
0.489	0.529 - $M_1 = 0.489$	$2L_2 + L_3 + M_1 = 0.487$
0.491	0.529 - $M_2 = 0.492$	$L_1 + L_2 + L_3 + M_2 = 0.491$
0.497	0.529 - $M_3 = 0.497$	$2L_1 + L_2 + M_2 = 0.497$
0.519	0.529 - $N_1 = 0.520$	$2L_1 + L_2 + M_2 = 0.520$
0.526	0.529 - O = 0.527	$2L_2 + L_3 + 2M_1 = 0.527$

There are a number of cases in which a set of several low energy beta rays have energies equal to  $E - K$ ,  $E - L_1$ ,  $E - L_2$ , and so on, where E is a constant which is usually supposed to be the energy of a gamma ray from the nucleus. As an example we will take the low energy beta rays from radium B, which have recently been remeasured very exactly by Ellis.\* These rays are supposed to indicate the presence of a nuclear gamma ray of energy 0.529. According to the present theory this gamma ray may be supposed to be emitted by the electronic system. Its energy is nearly equal to  $L_1 + L_2 + L_3 + 2M_2 = 0.528$ , or to  $K - L_1 - L_2 - M_1 - M_2 = 0.529$ . Table IV gives two alternative representations of these beta rays. The first one is the

\* 'Proc. Roy. Soc.,' A, vol. 143, p. 350 (1934).

same as that given by Ellis who regarded the ray of energy 0.529 as a nuclear gamma ray.

The second representation has the advantage that it takes in the two rays 0.444 and 0.469. It is interesting in that it includes nearly all the possible combinations of the three L levels.

These cases are sufficient to show that the low energy beta rays may be regarded as emissions by the electronic system after it has been excited. No absolute proof of this can be given at present, but it seems to be a reasonable supposition.

The higher energy rays will now be considered. According to the theory proposed the total energy of an atom should be equal to  $nq + \sum N_m E_m$  where  $q = 3.85$ . The term  $nq$  may be said to be the energy of the nucleus and  $\sum N_m E_m$  that of the electrons. Strictly speaking this hypothetical division of the energy between the nucleus and the electrons is not justified, but it is convenient and it facilitates a description of what may be imagined to take place.

We suppose that transitions are possible from one possible state to another of lower energy the difference being emitted as a gamma or beta ray. If  $n$  and  $N_m$  are the values before the emission and  $n'$  and  $N_m'$  those after, the energy of the ray emitted is  $(n - n')q + \sum (N_m - N_m') E_m$ . Putting  $n - n' = n''$  and  $N_m - N_m' = N_m''$  and then dropping the double dashes, we get for the energy of the ray emitted  $nq + \sum N_m E_m$ . In this last expression the  $N_m$ 's may be either positive or negative integers.

Suppose, for example, that a gamma ray is emitted with energy  $nq - K - 2L_1 - M_1$ . This can be supposed to be due to a gamma ray of energy  $nq$  emitted by the nucleus which raises four electrons to zero levels and then escapes, leaving the electronic system with energy  $K + 2L_1 + M_1$ . If the electrons or some of them are ejected from the atom we might get a gamma ray with less energy than  $nq - K - 2L_1 - M_1$ . Such ionization might result in a continuous energy distribution for the emitted rays and it may perhaps sometimes occur. The total energy emitted must, however, be  $nq + \sum N_m E_m$  and it seems that frequently, if not always, a single gamma or beta ray is emitted with this energy.

Corresponding to a ray with, for example, energy  $nq - K - 2L_1 - M_1$  we might expect to have one with energy  $n'q + K + 2L_1 + M_1$ . Where the nucleus emits several gamma rays during one disintegration, if one ray leaves the electrons with energy  $K + 2L_1 + M_1$  the next one could pick up this energy from the electrons and so come off with energy  $n'q + K + 2L_1 + M_1$ .

The gamma and beta ray energies from radium C' form a great many pairs with sums which are multiples of  $q$ . Such pairs on the present theory are due to rays with energies  $n'q + \sum N_m E_m$  and  $nq - \sum N_m E_m$  the  $N_m$ 's having the same values in both expressions.

The normal radium C' disintegrations have energy  $78.29$  or  $20q + 1.29$ , but a few occur with energies up to  $107.09$  or  $28q - 0.71$ , so that, presumably, in the normal disintegrations the difference  $8q$  goes into gamma rays. There should, therefore, be several gamma rays with total energy  $8q$ . The most intense gamma ray has energy  $6.07 = 2q - 1.63$ .

Table V gives a number of such pairs for radium C'. Gamma and beta ray energies were both used in forming these pairs. The "gamma ray" energies are merely the beta ray energies plus  $K = 0.9301$ .

Table V.

$26.06 + 20.16 = 46.22$	$13.36 + 13.59 = 26.95$
$21.21 + 25.00 = 46.21$	$5.90 + 21.04 = 26.94$
$22.64 + 23.57 = 46.21$	$9.91 + 17.05 = 26.96$
$22.14 + 24.07 = 46.21$	$9.34 + 17.60 = 26.94$
$25.13 + 21.09 = 46.22$	$13.15 + 13.79 = 26.94$
$3.85 \times 12 = 46.20$	$5.40 + 21.51 = 26.91$
	$14.11 + 12.86 = 26.97$
$21.21 + 17.29 = 38.50$	$5.14 + 21.82 = 26.96$
$17.46 + 21.04 = 38.50$	$16.67 + 10.28 = 26.95$
$14.16 + 24.32 = 38.48$	$8.98 + 17.98 = 26.96$
$21.97 + 16.53 = 38.50$	$20.44 + 6.50 = 26.94$
$14.42 + 24.07 = 38.49$	$4.98 + 21.97 = 26.95$
$22.14 + 16.36 = 38.50$	$10.63 + 16.36 = 26.99$
$18.39 + 20.11 = 38.50$	$4.59 + 22.35 = 26.94$
$22.35 + 16.17 = 38.52$	$14.87 + 12.09 = 26.96$
$21.82 + 16.67 = 38.49$	$11.97 + 14.99 = 26.96$
$13.36 + 25.13 = 38.49$	$11.16 + 15.78 = 26.94$
$20.58 + 17.89 = 38.47$	$3.66 + 23.28 = 26.94$
$20.16 + 18.35 = 38.51$	$11.45 + 15.53 = 26.98$
$3.85 \times 10 = 38.50$	$7.66 + 19.29 = 26.95$
	$11.56 + 15.37 = 26.93$
	$3.85 \times 7 = 26.95$
$21.21 + 5.72 = 26.93$	
$21.09 + 5.84 = 26.93$	
$25.13 + 1.79 = 26.92$	

Similar sets can be found with sums equal to any multiple of  $q$  less than 12. The emission of gamma rays with energies  $nq \pm C$  may be compared with the Raman effect in which a photon either gains energy from an atom or gives energy to it.

There are 67 beta ray energies for radium C' and 67 "gamma ray" energies. These give 134 values of  $C$  in  $nq \pm C$  which all lie between 0 and  $1.925$ . The  $C$ 's are, of course, all nearly equal to values of  $\sum N_m E_m$ , so that just as with the low energy beta rays, support for the theory that  $C = \sum N_m E_m$  can only be

obtained by looking for adjacent C's which are equal to very simple cases of  $\Sigma N_m E_m$ .

In a list of the C's in order of size there are two nearly equal to zero and then four values, viz., 0.028, 0.034, 0.036, and 0.049, which are nearly equal to one or other of the three M levels,  $M_3 = 0.033$ ,  $M_2 = 0.038$ , and  $M_1 = 0.041$ . Next there are two values 0.075 and 0.101 which are nearly equal to  $2M_2 = 0.076$  and  $3M_3 = 0.099$ . There are then seven values nearly equal to an L level, or to an L level plus an M level, viz., 0.131, 0.140, 0.179, 0.184, 0.194, 0.198, and 0.198. The next value is 0.262 which is nearly equal to

Table VI.

Gamma, beta, or alpha ray energies.	$\Sigma N_m E_m$ .
$q + 0.928$	K = 0.930
$3q - 0.922$	K - $M_1 + M_1 = 0.922$
$23q - 0.920$	K - $N_1 = 0.920$
$q - 0.916$	
$5q + 0.910$	K - $L_2 + L_3 = 0.906$
$4q + 0.902$	K - $L_1 + L_3 = 0.899$
$5q - 0.896$	K - $M_2 = 0.897$
$7q - 0.888$	K - $M_2 = 0.892$
$4q + 0.882$	K - $M_1 = 0.889$
$2q - 0.871$	
$26q + 0.870$	K - $M_1 - N_1 = 0.879$
$5q - 0.864$	K - $2M_3 = 0.864$
$5q + 0.855$	K - $2M_2 = 0.854$
$3q + 0.850$	K - $2M_1 = 0.848$
$3q + 0.829$	K - $3M_3 = 0.831$
$4q - 0.799$	K - $L_3 = 0.792$
$2q - 0.776$	
$4q + 0.765$	K - $L_3 = 0.768$
$q - 0.751$	K - $L_3 - M_1 = 0.751$
$6q - 0.746$	
$q + 0.736$	K - $L_2 - M_2 = 0.735$
$2q - 0.725$	K - $L_1 - M_1 = 0.720$

$L_1 + 2M_1 = 0.251$ . These small values of the C's are thus consistent with the theory that  $C = \Sigma N_m E_m$ .

Table VI gives the C's between 0.928 and 0.725. Some disintegration energy C's are included.

Table VII gives the C's between 0.549 and 0.508.

The C's in this set are very similar to the low energy beta rays of radium B in Table IV. The agreement between these C's and the combinations of three L levels and one M level is probably not without significance.

Table VIII gives the C's between 1.90 and 1.695.

The simple values of  $\Sigma N_m E_m$  in this set appear to indicate that the equation  $C = \Sigma N_m E_m$  may have real significance.

Table VII.

Gamma, beta, and alpha ray energies.	$\Sigma N_m E_m$ .
$q - 0.549$	$2L_1 + L_2 + M_1 = 0.541$
$3q + 0.537$	$L_1 + 2L_2 + M_1 = 0.535$
$4q + 0.531$	$L_1 + 2L_2 + M_2 = 0.531$
$27q - 0.530$	
$4q + 0.522$	$L_1 + 2L_2 + M_3 = 0.526$
$q - 0.516$	$2L_1 + L_2 + M_1 = 0.517$
$5q - 0.513$	$2L_1 + L_2 + M_2 = 0.514$
$3q - 0.510$	
$16q - 0.510$	$L_1 + L_2 + L_3 + M_1 = 0.510$
$4q - 0.508$	$2L_1 + L_2 + M_3 = 0.508$

Table VIII.

Gamma, beta, and alpha ray energies.	$\Sigma N_m E_m$ .
$14q + 1.90$	
$6q + 1.897$	$2K + M_1 = 1.901$
$6q - 1.892$	$2K + M_2 = 1.893$
$4q + 1.885$	
$q + 1.870$	
$1.866$	
$2q - 1.864$	$2K = 1.860$
$q + 1.858$	
$5q + 1.840$	$2K - M_2 = 1.827$
$7q - 1.818$	
$4q - 1.812$	$2K - M_1 = 1.819$
$3q + 1.805$	
$2q - 1.801$	
$5q - 1.794$	$2K - 2M_2 = 1.794$
$5q + 1.793$	
$1.791$	$2K - 2M_2 = 1.784$
$3q + 1.780$	$2K - 2M_1 = 1.778$
$q + 1.717$	$2K - L_2 = 1.722$
$2q - 1.705$	$2K - L_2 = 1.698$
$4q + 1.695$	$2K - L_1 = 1.691$

There are a considerable number of cases in which the difference between two C's is equal to K not including, of course, C's for a beta ray, and the gamma ray got by adding K to the beta ray energy. A number of examples are given in Table IX.

The C's between 0.414 and 0.316 are all equal to combinations like  $L + L' + M$  or  $L + L' + M + M'$ , where L and M stand for any L or M



level, so that the C's greater by K are equal to  $K + L + L' + M$ , or  $K + L + L' + M + M'$ .

If the C's for radium C' are really equal to  $\Sigma N_m E_m$  then we should expect to be able to represent the C's for the beta and gamma ray energies of the other elements in the same way. This may be examined by selecting a value of  $\Sigma N_m E_m$  such as  $K + 2L_1 + M_1$  and then seeing if values equal to

Table IX.

Ray energies.		$C_1 - C_2$
$nq \pm C_1$	$nq \pm C_2$	
+ 1.343	$q + 0.414$	0.929
$5q + 1.328$	$4q + 0.399$	0.929
$25q - 1.320$	$28q - 0.390$	0.930
$3q + 1.309$	$q - 0.377$	0.932
$20q + 1.290$	$25q + 0.360$	0.930
$24q + 1.280$	$6q - 0.348$	0.932
$6q - 1.278$	$3q - 0.345$	0.933
$27q - 1.260$	$18q - 0.330$	0.930
$3q - 1.246$	$2q - 0.316$	0.930

$nq \pm (K + 2L_1 + M_1)$  appear among the beta and gamma energies of other radioactive elements.

This method is a better test of the theory that  $E = nq + \Sigma N_m E_m$ , where E denotes a beta or gamma ray energy, because we select a particular value of  $\Sigma N_m E_m$  and see if it is present instead of finding values of  $\Sigma N_m E_m$  equal to observed values of  $E - nq$  which can usually be done. Suppose, for example, we select a value of  $\Sigma N_m E_m$  and then find it with, say, five or six out of the seven or eight elements which have more than a very few beta rays. It is clear that the chance of this being due to accidental numerical coincidences is small. Table X gives a number of such examples. Beta and gamma rays of low energy are included in this table because most elements have very few rays of high energy.

The results in Table X seem to make it probable that the expression  $nq + \Sigma N_m E_m$  has real significance when used to represent the energies of beta and gamma rays.

When the atom disintegrates with the emission of an alpha ray we assume that the disintegration energy must be equal to the difference between a possible energy before the disintegration and another possible energy after the disintegration. Just as when gamma or beta rays are emitted the whole

Table X.

Element.	$\Sigma N_m E_m$ .	$E - nq$ (observed).
Mesothorium 2	$K + 2L_1 + 4L_2 + M_1 = 2.200$	2.210 $\beta$
Radioactinium	$K + 2L_1 + 4L_2 + M_1 = 2.083$	2.097 $\beta$
Radium C'	$K + 2L_1 + 4L_2 + M_1 = 1.861$	1.862 $\gamma$
Radium C'	$K + 2L_1 + 4L_2 + M_1 = 1.861$	1.866 $\beta$
Radium B	$K + 2L_1 + 4L_2 + M_1 = 1.805$ ( $K + 2L_1 + 4L_2 + M_1 = 2K$ nearly)	1.806 $\gamma$
Mesothorium 2	$K + 2L_1 + 2L_2 + L_3 + M_1 = 2.106$	2.099 $\beta$
Actinium X	$K + 2L_1 + 2L_2 + L_3 + M_1 = 1.877$	1.870 $\gamma$
Radium C'	$K + 2L_1 + 2L_2 + L_3 + M_1 = 1.771$	1.780 $\gamma$
Radium B	$K + 2L_1 + 2L_2 + L_3 + M_1 = 1.717$	1.694 $\gamma$
Thorium C''	$K + 2L_1 + 2L_2 + L_3 + M_1 = 1.666$	1.640 $\gamma$
Protactinium	$K + 2L_1 + 2L_2 + M_1 = 1.802$	1.896 $\beta$
Radioactinium	$K + 2L_1 + 2L_2 + M_1 = 1.837$	1.826 $\gamma$
Radium C'	$K + 2L_1 + 2L_2 + M_1 = 1.633$	1.632 $\gamma$
Thorium B	$K + 2L_1 + 2L_2 + M_1 = 1.583$	1.593 $\beta$
Thorium C. C''	$K + 2L_1 + 2L_2 + M_1 = 1.488$	1.473 $\beta$
Radioactinium	$K + L_1 + 3L_2 + M_1 = 1.737$	1.729 $\beta$
Radium C'	$K + L_1 + 3L_2 + M_1 = 1.554$	1.554 $\gamma$
Radium C'	$K + L_1 + 3L_2 + M_1 = 1.554$	1.558 $\beta$
Radium B	$K + L_1 + 3L_2 + M_1 = 1.508$	1.504 $\beta$
Mesothorium 2	$K + 2L_1 + M_1 = 1.552$	1.559 $\beta$
Radioactinium	$K + 2L_1 + M_1 = 1.467$	1.460 $\gamma$
Actinium X	$K + 2L_1 + M_1 = 1.385$	1.387 $\beta$
Radium C'	$K + 2L_1 + M_1 = 1.310$	1.314 $\gamma$
Radium C'	$K + 2L_1 + M_1 = 1.310$	1.309 $\beta$
Radium B	$K + 2L_1 + M_1 = 1.270$	1.279 $\gamma$
Thorium B	$K + 2L_1 + M_1 = 1.270$	1.270 $\beta$
Thorium C''	$K + 2L_1 + M_1 = 1.229$	1.241 $\gamma$
Mesothorium 2	$K + L_1 + 2M_1 = 1.400$	1.406 $\beta$
Radioactinium	$K + L_1 + 2M_1 = 1.323$	1.313 $\gamma$
Actinium X	$K + L_1 + 2M_1 = 1.250$	1.259 $\beta$
Radium C''	$K + L_1 + 2M_1 = 1.182$	1.192 $\gamma$
Radium B	$K + L_1 + 2M_1 = 1.146$	1.141 $\beta$
Thorium B	$K + L_1 + 2M_1 = 1.146$	1.147 $\gamma$
Mesothorium 2	$K + L_1 + M_1 = 1.348$	1.337 $\beta$
Actinium X	$K + L_1 + M_1 = 1.305$	1.217 $\gamma$
Radium C'	$K + L_1 + M_1 = 1.140$	1.129 $\gamma$
Radium B	$K + L_1 + M_1 = 1.106$	1.118 $\gamma$
Thorium B	$K + L_1 + M_1 = 1.106$	1.107 $\beta$
Mesothorium 2	$K + L_2 - M_1 = 1.237$	1.245 $\beta$
Radioactinium	$K + L_2 - M_1 = 1.172$	1.176 $\gamma$
Actinium X	$K + L_2 - M_1 = 1.108$	1.111 $\gamma$
Radium C'	$K + L_2 - M_1 = 1.051$	1.050 $\gamma$
Radium C'	$K + L_2 - M_1 = 1.051$	1.056 $\beta$
Thorium B	$K + L_2 - M_1 = 1.020$	1.014 $\gamma$
Mesothorium 2	$K + M_2 = 1.131$	1.129 $\beta$
Protactinium	$K + M_2 = 1.103$	1.104 $\beta$
Radioactinium	$K + M_2 = 1.070$	1.074 $\beta$
Radium C'	$K + M_2 = 0.962$	0.962 $\gamma$
Radium C'	$K + M_2 = 0.962$	0.966 $\beta$
Radium B	$K + M_2 = 0.935$	0.928 $\gamma$

Table X—(continued).

Element.	$\Sigma N_m E_m$ .	E — $nq$ (observed).
Mesothorium 2	K = 1.092	1.093 $\beta$
Radioactinium	K = 1.035	1.029 $\beta$
Radium C'	K = 0.930	0.928 $\beta$
Radium B	K = 0.903	0.913 $\gamma$
Thorium C''	K = 0.875	0.870 $\gamma$
Mesothorium 2	K — M <sub>1</sub> = 1.040	1.062 $\gamma$
Radioactinium	K — M <sub>1</sub> = 0.987	0.990 $\beta$
Radium C'	K — M <sub>1</sub> = 0.889	0.888 $\gamma$
Thorium B	K — M <sub>1</sub> = 0.863	0.860 $\gamma$
Radium B	K — M <sub>1</sub> = 0.803	0.858 $\beta$
Mesothorium 2	K — M <sub>2</sub> = 1.051	1.062 $\gamma$
Radioactinium	K — M <sub>2</sub> = 1.000	1.007 $\gamma$
Radium C'	K — M <sub>2</sub> = 0.898	0.902 $\gamma$
Radium C''	K — M <sub>2</sub> = 0.898	0.896 $\beta$
Thorium C''	K — M <sub>2</sub> = 0.845	0.850 $\gamma$
Mesothorium 2	K — 2L <sub>2</sub> — M <sub>2</sub> = 0.733	0.730 $\beta$
Radioactinium	K — 2L <sub>2</sub> — M <sub>2</sub> = 0.692	0.689 $\gamma$
Radium C'	K — 2L <sub>2</sub> — M <sub>2</sub> = 0.622	0.614 $\gamma$
Radium C''	K — 2L <sub>2</sub> — M <sub>2</sub> = 0.622	0.624 $\beta$
Radium B	K — 2L <sub>2</sub> — M <sub>2</sub> = 0.604	0.607 $\gamma$
Thorium C. C''	K — 2L <sub>2</sub> — M <sub>2</sub> = 0.572	0.595 $\beta$
Mesothorium 2	K — 2L <sub>1</sub> — M <sub>1</sub> = 0.632	0.631 $\beta$
Radioactinium	K — 2L <sub>1</sub> — M <sub>1</sub> = 0.603	0.598 $\beta$
Radium C'	K — 2L <sub>1</sub> — M <sub>1</sub> = 0.551	0.549 $\gamma$
Radium B	K — 2L <sub>1</sub> — M <sub>1</sub> = 0.536	0.529 $\gamma$
Mesothorium 2	L <sub>1</sub> + L <sub>2</sub> + L <sub>3</sub> = 0.562	0.566 $\beta$
Radioactinium	L <sub>1</sub> + L <sub>2</sub> + L <sub>3</sub> = 0.531	0.532 $\gamma$
Actinium X	L <sub>1</sub> + L <sub>2</sub> + L <sub>3</sub> = 0.499	0.500 $\gamma$
Radium C'	L <sub>1</sub> + L <sub>2</sub> + L <sub>3</sub> = 0.469	0.467 $\gamma$
Radium C''	L <sub>1</sub> + L <sub>2</sub> + L <sub>3</sub> = 0.469	0.463 $\beta$
Radium B	L <sub>1</sub> + L <sub>2</sub> + L <sub>3</sub> = 0.454	0.444 $\beta$
Mesothorium 2	2L <sub>1</sub> + M <sub>1</sub> = 0.460	0.467 $\gamma$
Radioactinium	2L <sub>1</sub> + M <sub>1</sub> = 0.432	0.437 $\gamma$
Radium C'	2L <sub>1</sub> + M <sub>1</sub> = 0.379	0.377 $\gamma$
Radium B	2L <sub>1</sub> + M <sub>1</sub> = 0.367	0.367 $\beta$

energy difference goes into a single ray so we may suppose that this is also true for alpha rays.

In the disintegration energies the atomic number is changed by the emission of the alpha ray so that the energy emitted should be

$$(n - n') q + \Sigma (N_m E_m - N'_m E'_m).$$

The disintegration energies are only known to about 0.05 so that the changes in the energy levels due to the disintegration are not important. If the electronic system is unexcited before the disintegration then  $N = 0$  so that the energy emitted is  $(n - n') q - \Sigma N'_m E'_m$ .

The relations between the gamma and beta ray energies and the disintegration energies according to this theory are very simple. Denoting a gamma or beta ray energy by  $E$  and a disintegration energy by  $E_d$  we have

$$E = nq + C$$

and

$$E_d = n'q + C'.$$

If  $C = C'$  then

$$E_d - E = (n' - n)q$$

and if  $C = -C'$  then

$$E_d + E = (n' + n)q.$$

Since the  $C$ 's for radium  $C'$  rays and the  $C$ 's for the disintegration energies are often equal,\* it follows that these relations are often satisfied, within the limits of error. This merely means, for example, that if we take the disintegration energy 78.29 of radium  $C'$  then since it is equal to  $20q + 1.29$  we can get all the gamma and beta ray energies equal to  $nq + 1.29$  by subtracting multiples of  $q$  from 78.29, and all the gamma and beta ray energies equal to  $nq - 1.29$  by subtracting 78.29 from multiples of  $q$ . In this way a number of gamma and beta ray energies are related to each disintegration energy. For example, there are about 10 gamma and beta ray energies of radium  $C'$  for which  $C$  is nearly equal to  $\pm 1.29$ , besides those of other elements.

According to Gamow and Rutherford's theory the difference between two disintegration energies for an atom is equal to a possible gamma ray energy for the same atom. According to the present theory the difference between any two disintegration energies

$$(n - n')q + \Sigma (N_m E_m - N'_m E'_m)$$

and

$$(n'' - n''')q + \Sigma (N''_m E''_m - N'''_m E'''_m)$$

is equal to

$$(n - n' - n'' + n''')q + \Sigma (N_m E_m - N'_m E'_m - N''_m E''_m + N'''_m E'''_m).$$

This is equal to  $nq + \Sigma N_m \bar{E}_m$ , where  $\bar{E}_m$  is a sort of average value of the energy levels  $E_m, E'_m, E''_m, E'''_m$ . Usually the differences between the energy levels of different atoms are not big enough to matter because the probable errors of the disintegration energies is 0.05 or more. The difference between any two disintegration energies should, therefore, be approximately equal to a possible beta or gamma ray energy for any atom.

\* Wilson, 'Phys. Rev.', vol. 44, p. 858, November 15 (1933).

There are about three hundred differences between the thirty-three disintegration energies,\* which lie within the range of observed gamma and beta ray energies. This range is from zero to about thirty so that the average number of disintegration differences per unit range is about ten. Thus if any number between zero and thirty is selected the most probable number of disintegration energy differences differing from it by not more than 0.05 is about one. This makes it difficult to test the theory that differences between disintegration energies are nearly equal to possible gamma and beta ray energies, because the probable error of the disintegration energies is 0.05 or more.

I find that there are a number of cases in which several disintegration energy differences do not differ from their mean value by more than 0.05. These differences which occur several times are often equal to prominent gamma or beta ray energies.

Table XI gives all the cases where three or more disintegration energy differences are equal within  $\pm 0.05$ , and the energies of gamma and beta

Table XI.

$E_d - E_d'$	Gamma and beta ray energies.
27.63 (3)	RaC' 27.4 $\beta$
22.70 (4)	RaC' 22.64 $\beta$
19.09 (3)	RaC' 19.18 $\beta$ , 19.28 $\gamma$
16.63 (4)	RaC' 16.67 $\beta$ , 16.53 $\gamma$
14.12 (4)	RaC' 14.11 $\beta$ , 14.20 $\gamma$
13.20 (3)	RaC' 13.23 $\beta$ , 13.15 $\gamma$
12.84 (3)	RaC' 12.86 $\beta$ , 12.80 $\gamma$
11.92 (3)	RaC' 11.87 $\beta$ , 11.97 $\gamma$
11.55 (4)	RaC' 11.45 $\beta$ , 11.56 $\gamma$
11.19 (3)	RaC' 11.16 $\beta$ , 11.20 $\gamma$
10.85 (3)	RaC' 10.63 $\beta$
10.52 (4)	RaC' 10.63 $\beta$
10.32 (3)	RaC' 10.28 $\beta$
9.82 (4)	RaC' 9.91 $\gamma$ , MesoTh2 9.70 $\gamma$
8.95 (3)	RaC' 8.98 $\beta$ , MesoTh2 8.97 $\beta$
8.54 (5)	ThB, etc., 8.55 $\beta$
8.47 (4)	RaC' 8.41 $\beta$
7.74 (4)	RaC' 7.66 $\gamma$ , ThB, etc., 7.72 $\beta$
7.30 (4)	RaC' 7.38 $\beta$ , ThB, etc., 7.22 $\beta$
6.83 (3)	RaC' 6.83 $\gamma$ , ThB, etc., 6.82 $\beta$
6.08 (7)	RaC' 6.05 $\beta$ , 6.07 $\gamma$
5.54 (5)	RaC' 5.57 $\beta$
5.45 (5)	RaC' 5.40 $\gamma$
5.05 (3)	RaC' 5.14 $\beta$ , ThC'' 5.07 $\gamma$
4.36 (3)	Ac(B + C) 4.40 $\beta$ , MesoTh2 4.42 $\gamma$
4.25 (4)	RaC' 4.26 $\gamma$ , ThB, etc., 4.22 $\beta$
3.70 (4)	RaC' 3.66 $\beta$ , RaB 3.76 $\beta$ , 3.78 $\gamma$ , ThB, etc., 3.66 $\beta$
2.90 (3)	RaC' 2.93 $\beta$ , RaB 2.88 $\beta$
2.73 (3)	RaC' 2.72 $\gamma$ , RaB 2.77 $\beta$ , ThC. C'' 2.72 $\beta$
2.14 (3)	ThC. C'' 2.13 $\beta$ , RaAc 2.12 $\beta$
1.82 (4)	RaC' 1.86 $\beta$ , RaB 1.83 $\beta$ , MesoTh2 1.84 $\gamma$ , AcX 1.80 $\beta$
1.75 (6)	RaC' 1.79 $\beta$ , MesoTh2 1.78 $\beta$ , RaAc 1.75 $\beta$ , AcX 1.76 $\beta$
1.33 (4)	RaC' 1.34 $\gamma$ , MesoTh2 1.33 $\beta$ , RaAc 1.31 $\beta$ , AcX 1.36 $\beta$ , ThB, etc., 1.27 $\beta$

\* Including the value 57.86 given by Briggs for Thorium-X. 'Proc. Roy. Soc.' A, vol. 143, p. 604 (1934).

rays which have nearly equal values. The first column gives the mean disintegration energy differences, and the number of which the mean was taken. Thus, for example, 6.08 (7) means that there are seven differences with values between 6.03 and 6.13. Differences greater than 28 or less than 1.30 were not considered.

Table XI gives 123 cases in which the difference between two disintegration energies is equal within the limits of error to beta and gamma ray energies. Often the gamma and beta rays are those of considerable intensity.

It is noteworthy that the only case in which there are seven disintegration energy differences nearly equal is for the value 6.08 which is the energy of the most intense gamma ray from radium C'. The gamma and beta ray energies are equal to  $nq + \sum N_m E_m$  and the disintegration energy differences are approximately equal to a similar expression. A probable value of  $\sum N_m E_m$  must be equal to the difference between two probable values. Thus we might expect the most probable values of the disintegration energy differences to agree with the most probable gamma and beta ray energies.

Ellis has pointed out a case in which the difference between two gamma ray energies is very exactly equal to the difference between two electronic energy levels. He comments on this as follows: "While there is no other case so well authenticated as this, there are indications that there are other gamma rays showing this curious relation that their energies differ by an amount exactly equal to energy differences of the electronic structure. It may be only a coincidence and certainly further investigation is needed before this observation would justify an attempt at a theoretical interpretation."

It is clear that such cases are to be expected on the present theory and that they help to confirm it.

Another case which has hitherto seemed anomalous is the intense beta ray of energy 13.23 from radium C', which is not accompanied by a gamma ray of energy  $13.23 + K$ . According to the present theory this beta ray is due to a nuclear gamma ray of energy  $3q = 11.55$  which receives energy 1.68 from the electrons. It may be supposed to follow the intense gamma ray of energy 6.07, which is due to a nuclear ray of energy  $2q = 7.70$ , which gives energy 1.63 to the electrons.

According to the present theory the chance that a nuclear gamma ray can escape from the atom without any interaction with the electrons must be regarded as very small, but it should be noted, as has already been pointed out, that the interaction may take place between the electrons and the nucleus before the nuclear gamma ray is emitted. The coupling between the electrons

and the nucleus is probably very important; in fact, it is obvious that the electronic energy levels are determined very largely by the nucleus rather than by the mutual actions between the electrons. Strictly speaking the present theory only requires the energy of the whole atom to be equal to  $nq + \sum N_m E_m$  and it is only for convenience that this energy is regarded as belonging partly to the nucleus and partly to the electrons. It is also assumed that when a transition to a state of lower energy takes place, the whole energy difference, usually at any rate, goes into a single ray which may be either a gamma, beta, or alpha ray. Apparent exceptions to this rule can be attributed to two or more transitions immediately following each other.

In conclusion, it may be said that there is a considerable body of evidence which appears to support the theory proposed so that it appears to be worthy of serious consideration.

I wish to thank Dr. C. D. Ellis, F.R.S., for very kindly letting me have his new very exact beta ray energy values before they were published.

#### *Summary.*

Reasons are given for believing that all low energy gamma and beta rays are emitted by the electronic system after it has been excited by a gamma ray from the nucleus. The higher energy beta and gamma ray energies are compared with the theory, previously suggested, that the nuclei of all the radioactive atoms emit rays with energies which are equal to multiples of  $3.85 \times 10^6$  electron volts. The disintegration energies are also considered on this theory. It is concluded that there is a considerable body of evidence which appears to support the theory.

---

*Note on a New Transition produced by Electron Impact in Helium.*

By R. WHIDDINGTON, F.R.S., and H. PRIESTLEY, B.Sc., The University of Leeds.

(Received February 27, 1934.)

[PLATE 9.]

During some experiments in this laboratory on electron impacts in helium, it was observed that in addition to electrons which had made the most probable transition impact ( $1\ ^1S_0-2\ ^1P_1$ ) there were a few which showed a loss indicating  $2\ (1\ ^1S_0-2\ ^1P_1)$  and others\* showing a loss approximately equal to  $3\ (1\ ^1S_0-2\ ^1P_1)$ .

The present note is concerned with this last mentioned loss.

The apparatus is of a type described before† in which the electrons under examination are deflected in a magnetic field and photographically recorded on an oiled photographic film. The electron spectra so obtained are very sharply defined even at low electron energies and permit of quite accurate measurement. An addition to the apparatus not previously described is a special film chamber which contains a cylindrical photographic film which can be rotated to any position from outside by a vacuum tight joint. As many as 24 photographs can now be taken on one film.

An idea of the type of photograph obtainable may be got from an inspection of fig. 1 (a), Plate 9, which is just a set of eight repetition spectra all taken at the same energy (200 volts) of impacting electrons and with the electron beam directed at an angle of  $3^\circ$  across the axis of the receiving slit system.

There are here to be seen clearly the line made by those elastically scattered electrons which have lost no energy (marked 0); the line made by those electrons which have lost energy in causing the transitions ( $2P$ ) once during their passage in the apparatus; the line made by those electrons which have lost twice this energy  $2\ (2P)$  at two separate inelastic impacts; and, finally, the faint line marked  $(2P)^2$  now under consideration.

The lines marked R are reference lines of known electron energies used for calibration purposes in measuring the film.

In the films on which measurements have actually been made, we have adopted the useful practice of repeating one of the reference lines in the clean

\* 'Nature,' vol. 131, p. 908 (1933); 'Proc. Leeds Phil. and Lit. Soc.,' vol. 2, p. 491 (1933).

† Jones and Whiddington, 'Phil. Mag.,' vol. 6, p. 890 (1928).



part of the spectrum between O and (2P). This ensures that there is always a definite point in the spectrum under measurement which may with certainty be regarded as corresponding exactly to the equivalent point in the series of reference lines.

The film reproduced does not show such a registration line—but in its place appears a faintish broad line. This is due to a molecular transition produced in a small amount of air in the apparatus and has been chosen in order to show the obvious width difference between lines due to atomic and molecular transitions.\*

The enlarged photograph in fig. 1 (b), Plate 9, shows the new (2P)<sup>2</sup> line particularly well.

It was thought at first that this line corresponded to a loss of energy  $3(2P) = 63.6$  volts—an interpretation, however, put forward with reserve in view of the relatively high intensity of the line and to the fact that only approximate measurements had been made on it.

Precise measurements have now been made.

It is not always possible to secure a photograph in which (2P)<sup>2</sup> is so well marked as to be regarded as really reliable for measurement purposes. Out of some four hundred photographs, twelve were selected as reaching the high standard required and, on measurement, gave the following mean results for the energy loss: 59.26, 59.17, 59.09, 59.16, 59.65, 58.91, 59.16, 59.17, 59.23, 59.55, 59.36, 59.31.

From these values—each of which represents the mean of a set of three separate determinations—the most probable error was calculated, giving the result:

$$59.25 \pm 0.12 \text{ volts.}$$

It is now quite clear that this loss does not correspond at all to any hitherto calculated transition.

There are two features of the photographs which deserve mention.

*The Width of the Line (2P)<sup>2</sup>.*—Although the width of this line is of the same order as that of a reference line, it appears visually to be somewhat broader. The general inference from this fact is that, whereas there is in all likelihood an excitation rather than an ionization process involved, yet there may be several such excitations differing only slightly in energy values.

Photometry, with a line of so low a density as this, is very difficult on account of the complicating influence of the background; but so far as we have been

\* Roberts and Whiddington, 'Proc. Leeds Phil. and Lit. Soc.', p. 46 (1930).

able to go in this direction there do seem to be indications of structure in the line which our apparatus is not quite competent to resolve.

*The Density of the Line.*—From the density of the line its relative probability can be inferred. It is roughly as dense as the line 2 (2P); that is, about the same number of electrons produce this new transition as produce two (2P) transitions in the apparatus at a pressure of 0.01 mm.

Thus if  $p$  is the probability of the 2P excitation, then  $p^2$  is the order of the probability of the new excitation—a result approximately true for the range of incident electron energy, 100–600 volts.

*The Nature of the Excitation.*—It seems pretty clear that the only likely process leading to the observed result is in the nature of excitation rather than ionization, as has already been pointed out.

A suggestion which seems possible is that this is an example of double excitation.

As the 2P state is the most probable single excitation, then a double excitation to this state involving both the S electrons would in all likelihood be also the most probable transition of this kind.

What the energy loss of this particular transition may be is not yet known, but we understand from Professor Hartree, of Manchester, that Professor Lindsay, of Providence, U.S.A., has just calculated on theoretical grounds the energy required in the double transition (2S)<sup>2</sup> and finds the value 58.6 volts. While this is not the double transition suggested by the present writers, it is not very different in value from that observed, and indeed, a similar calculation for the double transition (2P)<sup>2</sup> may perhaps show even closer accord with the experimental value.

Some of the initial work in these experiments was carried out by Mr. T. Emmerson, while in receipt of a grant from the Department of Scientific and Industrial Research. Our thanks are here recorded.

We have also to thank the Government Grant Committee of the Royal Society for a grant to the senior author, out of which was purchased the photometer.

#### *Summary.*

Electrons of energy between 100 and 600 volts, passing through helium at low pressure, have been observed to produce a new excitation transition with associated energy  $59.25 \pm 0.12$  volts. A rough estimate of the probability of the transition is made. It is suggested that it may be due to a double excitation of the helium atom involving the transition of both  $s$  electrons to  $p$  states.

---



Photometric curves showing, in helium, argon, and neon, the electron energy distribution in an initially homogeneous beam after elastic and inelastic encounters respectively.

*Small Angle Inelastic Electron Scattering in Helium, Neon, and Argon.*

By R. WHIDDINGTON, F.R.S., and J. E. TAYLOR, Ph.D., B.Sc., The University of Leeds.

(Received February 27, 1934.)

[PLATE 10.]

*Introduction.*

A considerable amount of work, both theoretical and experimental, in the domain of electron scattering in gases has been carried out in recent years, and a comprehensive and most useful treatise has recently been written by Mott and Massey on this subject.

Theory and experiment are in general relative agreement, it appears, at scattering angles greater than  $10^\circ$  and over wide ranges of electron energy, but so far attention has not been directed to the phenomena at angles much less than  $10^\circ$ —nor has an absolute determination of the probability of any type of collision been made experimentally. The present work is an attempt to obtain results in these directions.

In this laboratory a special photographic method has been developed for examining these effects which appears to have some special advantages of its own, at any rate at electron energies between 50 and 300 volts, the range studied.

In the particular work here described, the main electron stream is directed from a hot filament towards a system of slits, traversing in their path the gas under examination at a low pressure. The slits receive a mixed stream containing mainly the original electrons, but a few which, travelling in the same direction, have suffered elastic collisions and a few which have suffered inelastic ones.

After passage through the slits, these electrons are made to traverse a suitably arranged magnetic field and are finally received on an oiled photographic film which, on development, records their presence at the appropriate points.

Clearly the electrons which have elastically collided will be indistinguishable from those of the main stream, but the inelastic groups will appear at definite positions on the film and from their positions the corresponding energy losses

can be obtained\* while from their photometric densities the numbers of electrons involved should be calculable.

It may be useful here to compare the sensitiveness of the magnetic photographic method with the electrostatic Faraday cylinder method in use by many experimenters.

This method makes use of an electrostatic field for sifting out the electrons of various energies, thereafter pouring a selected part into a Faraday cylinder and measuring the charge by an electrometer or similar device. Such an arrangement responds to a current of about  $10^{-14}$  ampere spread over the area of the cylinder, which is often quite large.

In the photographic method a current of less than  $10^{-11}$  ampere per square millimetre can be measured with exposures of the order of 10 minutes. When it is pointed out that accurate photometry is feasible with areas less than 0.04 sq. mm. the current beam is seen to be only  $4 \times 10^{-13}$  ampere, which is at any rate within reach of the electrostatic method. There is, moreover, the immense advantage that on one and the same photograph are recorded electrons of all energies—elastic or inelastic—and they are strictly comparable in a relative way, since if there be variation of one, the other will vary in a comparative manner. It must be stressed that the electrostatic method is more sensitive so far as measurement of the quantities is concerned, but the other advantage to which reference has been made, in the opinion of the writers, is of overpowering value.

In this paper the probabilities of the most probable excitations in helium, argon, and neon are compared over the range of incident electron energies 50–350 volts, and for the particular case where the electron is apparently undeviated by the collision.

The helium excitation involved occurs with the transition ( $1^1S_0-2^1P_1$ ) which is equivalent spectroscopically to an energy change of 21.11 volts.†

Roberts and Whiddington showed‡ that this obtained for electron impact, although the energy loss suffered by the impacting electron was slightly higher at 21.24 volts. This work was later supported by the experiments of Van Atta.§

With argon the matter is not quite as simple, since there are two excitations differing so little in energy as to be unresolved by the apparatus.

\* Roberts and Whiddington, 'Phil. Mag.', vol. 12, p. 962 (1931).

† "Atomic Energy Levels," by Becher and Goudsmit.

‡ 'Phil. Mag.', vol. 12, p. 962 (1931).

§ 'Phys. Rev.', vol. 38, p. 876 (1931).

These excitations correspond to the transitions  $3^1S_0-4^3P_1$  (11.57 volts) and  $3^1S_0-4^1P_1$  (11.77 volts) the numbers in brackets being the associated energy changes calculated from spectroscopic data.

There seems no doubt that a combination of these transitions is the most probable over the range of incident electron energies considered, particularly as the work of Roberts and Whiddington\* has been confirmed by Van Atta.†

It is of interest, in passing, to note that Bullard and Massey‡ in their experiments failed to detect the presence of these transitions at an incident electron energy of 30 volts.

The most probable transitions in neon appear to be  $2^1S_0-3^3P_1$  (16.59 volts) and  $2^1S_0-3^1P_1$  (16.76 volts)—again not resolved in our apparatus.

### *The Apparatus.*

The general arrangement of apparatus has already been fully described elsewhere§ with only slight modifications. These modifications are included in fig. 1 which shows the anode A with its circular aperture 1 mm. in diameter

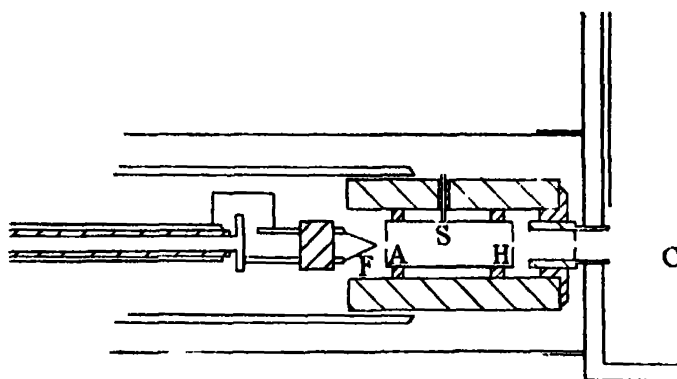


FIG. 1.—Diagram of apparatus (2 × size).

the emitting filament with its tip 1.25 mm. away from the anode and centrally fixed, and the collision space S with a small hole H only 0.145 mm. in diameter.

This hole not only limits the beam but also allows a gas pressure to build up in S, which is considerably in excess of the low pressure maintained in the camera space C.

\* 'Proc. Leeds Phil. and Lit. Soc.,' vol. 2, p. 46 (1930).

† *Loc. cit.*

‡ 'Proc. Roy. Soc.,' A, vol. 130, p. 579 (1931).

§ Roberts and Whiddington, 'Phil. Mag.,' vol. 38, p. 876 (1931).

Between C and H, 8.5 mm. in front of H, there is in addition a wide slit as shown in fig. 1.

All these parts are well shielded magnetically in thick iron tubes, so that the magnetic field which deflects the electrons in the camera is inoperative in the collision space.

Typical examples of photographs obtained have already been published,\* but the problem now is rather more difficult than the precise measurement of the positions of the loss lines—it involves the measurement of the photographic density and its expression in terms of electron intensity.

The matter is to some extent simplified by the fact that of the original electron stream only a very few suffer collisions with the gas atoms—a state of affairs consequent on a suitable adjustment of the gas pressure in the collision space.

The developed photograph thus shows two dark patches—one very dense and due to the main electron stream together with that small number of electrons which have suffered elastic collisions—the other relatively very faint and due to electrons which have suffered the proper energy loss corresponding to the most probable transition.

These are termed the full velocity line and the loss line respectively.

The evaluation of the electron intensities in these lines is carried out with the help of the photographic law which has been shown by the writers to hold.†

This law, which is similar to that for light, takes the form

$$D = \gamma \log It^n - i,$$

where  $D$  is the photographic density of the blackened point above that of its immediate background.  $I$  is the intensity of the electron beam which has produced the blackening for an exposure time  $t$ . Of the photographic constants,  $\gamma$  is particularly important in that its value is determined largely by development conditions.

This means, of course, that  $\gamma$  varies slightly from photograph to photograph even under the most carefully regulated conditions, so that it is not possible to obtain accurate results from different films unless each is separately calibrated.

This is done in a simple manner first by taking the long exposure (10 minutes, say) necessary to bring out the “loss line”—and then (by altering slightly

\* Roberts and Whiddington, ‘*Phil. Mag.*,’ vol. 12, p. 962 (1931).

† ‘*Proc. Roy. Soc.*,’ A, vol. 136, p. 651 (1932).



the deflecting magnetic field) putting on several "reference lines" of very short exposure\* so that the full velocity lines only come out in the photograph with densities comparable with and "straddling" the loss line.

From a series of such reference lines a  $(D, \log t)$  curve can be constructed, and, by intrapolation, the time of exposure ( $t_1$ ) of the main beam corresponding to the density of the loss line, exposed for time ( $t_2$ ) can be obtained.

Then

$$I_1 t_1^p = I_2 t_2^p$$

and

$$I_2/I_1 = (t_1/t_2)^p,$$

where  $I_1$  and  $I_2$  are the electron intensities of the main and loss line respectively.

For electrons,  $p$  has the value 0.88, so that  $I_2/I_1$  comes out readily from this relation.

A point which should be mentioned in connection with this photographic electron recording, is the fact discovered in this laboratory in 1925, that there is an abrupt change in the photographic blackening law in the neighbourhood of 60 electron volts. The law, however, is independent of electron energy up to about 400 volts. It is always necessary, therefore, when working with the lower energy electrons to accelerate them to a safe working speed after traversing the collision space. This is easily done and, of course, does not affect the relative numbers of "full velocity" and "loss" electrons in any way.

The density measurements were made from wedge readings on a Cambridge recording photometer.

The results are shown in graphical form in fig. 2 and Plate 10.

*Helium.*—This gas was admitted into the collision space so that the continuous flow maintained a pressure of 0.0120 mm. of mercury. At this pressure there were no signs of multiple collisions such as have been recorded by us elsewhere,† so that only single collisions were of importance.

The relative probabilities ( $I_2/I_1$ ) obtained in the manner already explained are reduced to a standard pressure of 0.0100 mm. for comparison purposes.

With helium the range of electron energy was varied between 40 and 400 volts, and over this range it is seen from the graph that the probability of excitation increases almost as the energy of the colliding electron.

\* A special time switch was constructed to effect these short exposures. It consisted of a heavy pendulum which operated knife switches so placed as to permit exposures varying in duration from about 0.05 second to 0.25 second. The switch was calibrated photographically by taking oscillograph records.

† 'Nature,' vol. 131, p. 908 (1933); 'Phil. Mag.,' vol. 6, p. 889 (1928).

The photometric curves of Plate 10 were taken across elastic and inelastic lines in the magnetic spectrum of helium and show clearly that the lines are of substantially equal width at equal ordinates measured from the respective

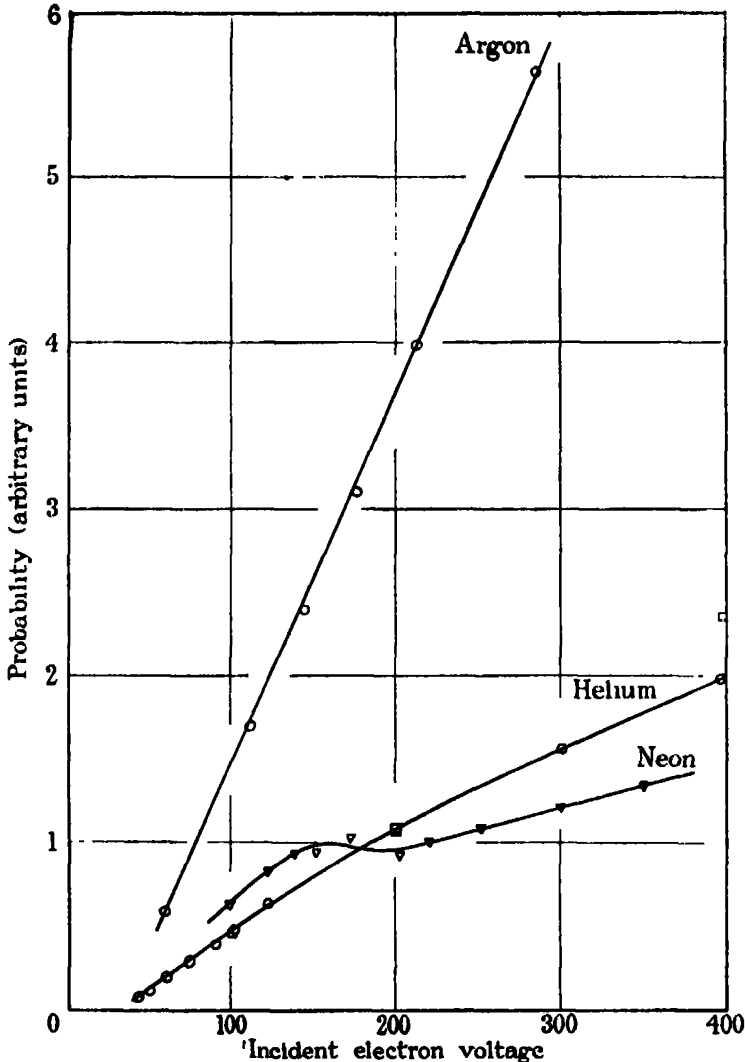


FIG. 2.—Curves showing variation of probability with incident electron energy for the most probable transitions in helium, argon and neon.  $\square$  = Probabilities for Helium calculated by Massey and Mohr.  $\circ$  = Experimental values.

peaks. This can only mean that no excitation other than ( $1^1S_0-2^1P_1$ ) is involved in the line considered.

*Argon.*—The measured probability for this gas is much higher than in helium at the same pressure, and the curve is a straight line within the limits of experimental error.

The photometric curve clearly shows that more than one excitation level may be involved in the measurements, since the inelastic curve is broader than the elastic one.

*Neon.*—The special feature of this curve is the pronounced maximum in the probability in the neighbourhood of 170 volts. The photometric curve for Neon again shows a broadened inelastic loss line so that, as with Argon, more than one excitation may be involved.

### Discussion of Results.

There are no theoretical calculations available, except for helium—and for this gas Massey and Mohr\* following Born, give data with which the present results can fortunately be easily compared.

No good results on absolute probability are so far available—though in this paper a preliminary attempt to make such a measurement is recorded—but the variation of the probability near zero angle with the electron energy shows fairly satisfactory agreement. In fig. 2 the calculated probability has been made to agree with our experimental value at 200 volts and it can be seen that over the range examined the values approximate closely to one another.

What divergence there is, might possibly be accounted for by allowing for the range of collision angles in the collision space, but any accurate computation along these lines has proved too baffling mainly on account of the impossibility of making reasonably proper assumptions in the neighbourhood of the emitting filament.

Table I.

Electron volts.	0°.	5°.	10°
100	7.8	4.4	1.78
200	17.7	4.5	0.99
400	39	2.6	0.33

For the singlet transition in helium, which alone concerns us, Table I taken from the paper by Massey and Mohr (*loc. cit.*, p. 618), shows the sort of variation of probability to be expected on theoretical grounds with angle of scattering.

\* 'Proc. Roy. Soc.,' A, vol. 140, p. 613 (1932).

The points plotted on our curve indicate a close adherence to the order of probabilities in the 0° column. Such divergence as our curve shows, is in the direction to be expected if a small number of electrons at the higher angles are included.

By inference from the better understood curve for holium, the argon curve is almost certainly due, in the main, to the transition ( $3^1S_0-4^1P_1$ ) corresponding to an energy loss in the colliding electron of 11.77 volts, but there is also a much less probable group of triplet transitions ( $3^1S_0-4^3P_{012}$ ) having values 11.49, 11.56, 11.66, of which ( $3^1S_0-4^3P_1$ ) is at the same time the most probable and about the average value 11.56 volts.

On account of the composite character of this line (a more probable 11.77 volts and a less probable 11.56 volts line) the observed line might be expected to be broader and with a value somewhat less than 11.77 volts.

These expectations are so far borne out by experiment. The photographic reduction of an argon photometric curve, Plate 10, shows clearly how the argon loss line is somewhat broader than the line produced by elastic collision. Further, the value of the main energy loss was measured by one of us (R. W.) and Roberts\* some years ago and found to be  $11.6 \pm 0.1$  volts, which is in qualitative agreement with the considerations just advanced.

With neon the position is not quite so clear cut, and it will be seen from the curves that there is a distinct maximum in the probability about 170 volts. The photometric curve in Plate 10 also makes it clear that, as with argon, there may be more than one transition involved.

The most probable transition seems to be the singlet ( $2^1S_0-3^1P_1$ ) with an energy change of 16.76 volts. There is also the triplet group ( $2^1S_0-3^3P_{012}$ ); the most probable member being ( $2^1S_0-3^3P_1$ ) with an energy of 16.58 volts—this is also about the average of the three.

As with argon, this results in a slight broadening of the line, and, although no measurements have yet been made, it seems very likely that the main ( $2^1S_0-3^1P_1$ ) loss of 16.76 volts will give experimentally a slightly lower value.

Similar experiments to those just described have been carried out by Van Atta† using an electrostatic measuring device instead of a photographic one.

When account is taken of the much larger angles included in his arrangement—there is general agreement with our results in helium and neon. Our curve for argon, however, is entirely different and we can see no way of explaining it.

\* 'Proc. Leeds Phil. and Lit. Soc.,' vol. 2, p. 46 (1930).

† 'Phys. Rev.,' vol. 38, p. 876 (1931).

We are, moreover, in disagreement with this author on one rather important point—we assign a relatively unimportant role to triplet transitions in our results.

*The Absolute Probability of the Excitation ( $1^1S_0 \rightarrow 2^1P_1$ ) by Electron Impact in Helium.*

The results of this section of the paper are not beyond criticism and are put forward with reserve. The subject is now being further investigated in this laboratory with apparatus of a different design, which, it is hoped, may lead to more reliable results. The electron beam used in these experiments carried 200-volt electrons.

*Method.*—It is obviously only necessary to compare the relative intensities of the full velocity line and the excitation loss line to arrive at an estimate of the probability. But to express the probability absolutely, requires a precise knowledge, not only of the volume within the collision space from which the electrons are drawn, but also of the intensity distribution within this space of the impacting electrons themselves.

Herein lies one of the more serious difficulties, since it may be nearest the filament that the greatest concentration occurs and it is precisely in this region that least is known, not only about this electron distribution, but also about the gas pressure.

One can, however, assume an inverse square law of distribution round the filament and integrate the whole effect from the camera pin hole to the anode hole—which is quite frankly only a rough approximation.

If this be done it appears that

$$\frac{I}{I_0} = \frac{\text{Intensity of loss line}}{\text{Intensity of full velocity line}} = \frac{l(l+d)}{(l-a)(l+d-a)} \cdot B \cdot N \cdot \pi r^2 \int \frac{dx}{x^2}.$$

In this formula  $B$  is the probability of excitation per unit solid angle (at 0.01 mm. of mercury) in the forward direction.  $a$  is the total path length traversed in the gas.  $d$  is the further distance travelled by electrons in the magnetic field.  $x$  is the distance of a slab of thickness  $dx$  from the filament.  $l$  is the distance measured from the filament, of the point of entry of the electrons into the magnetic field, so that  $(d+l)$  is the total path length of electrons coming from the filament.  $N$  is the number of atoms per cubic centimetre of the helium at the pressure used.  $r$  is the radius of the camera pin hole.

The scattered electrons above considered would, in the absence of a magnetic field, form a circular path on any screen held at right angles to their cone

axis (of semi-angle  $6^\circ$ ). But bent round in the magnetic field through  $2\pi$  it can be shown that an elliptical patch is formed.\*

In the above formula a "spreading factor" is included to take this into account.

When the integration is carried out between the limits set by the anode hole (A) and the camera pin hole (H), B expressed in  $\pi a_0^2$  units for a gas pressure of 1 mm. of mercury, is

$$2\pi B = 68.5,$$

which compares with 17.7 according to the theory set out by Massey and Mohr.

Had the integration been extended into the doubtful region of the filament, the result could have been made equal to the predicted one.

The photometric part of this investigation was carried out on a Cambridge recording photometer, purchased with a grant made to the senior author by the Government Grant Committee of the Royal Society. The use of this instrument has enormously reduced the labour necessary.

### *Summary.*

Experiments have been carried out to measure the relative probabilities of the most probable excitations by electron impact in helium, argon and neon.

The method is substantially the same as that already described elsewhere and consists of photographing the complete magnetic spectrum on films specially sensitized by an oiling process.

Only one excitation ( $1^1S_0-2^1P_1$ ) is involved in helium; two at least in argon—singlet and triplet ( $3^1S_0-4^1P_1$ ) and ( $3^3S_0-4^3P_1$ ); two at least in neon—singlet and triplet ( $2^1S_0-3^1P_1$ ) and ( $2^3S_0-3^3P_1$ ).

The presence of these various transitions is inferred from photographs which were afterwards photometered—some included in the text of the paper.

The variation of probability with electron energy at angles of scattering in the region of  $0^\circ$  between 50 and 400 volts agrees fairly well with the theory of Massey and Mohr for helium. Argon has a similar law of variation—nearly linear with electron energy—but neon has a maximum in the region of 170 volts, which has not been explained.

A very approximate estimate of the absolute probability for helium is given and is compared with the computation on theoretical grounds. The agreement is not unsatisfactory in view of the assumptions necessarily made—an experimental value of 68.5 comparing with a computed one of 17.7. This part of the work is not regarded as reliable and is being continued with improved apparatus.

\* Barber, 'Proc. Leeds Phil. and Lit. Soc.', vol. 2, p. 427 (1933).

## *The Viscosity of Strong Electrolytes Measured by a Differential Method.*

By W. M. Cox, B.A., B.Sc., Exhibitioner of Exeter College, and  
J. H. WOLFENDEN, M.A., Fellow of Exeter College, Oxford.

(Communicated by Sir Harold Hartley, F.R.S.—Received August 4, 1933.

Revised March 26, 1934.)

Until quite recently no satisfactory equation had been obtained for the representation of the viscosity of dilute solutions of strong electrolytes. An empirical equation was recently proposed by Jones and Dole\* to fit the only accurate data then available. Their equation may be represented thus :

$$\eta = 1 + A\sqrt{c} + Bc,$$

$\eta$  = relative viscosity of the solution

$c$  = concentration in moles per litre

$A$  and  $B$  are constants.

Jones and Dole realized that the coefficient  $A$  is due to interionic forces and in a series of later publications Falkenhagen, Dole and Vernon† have deduced a theoretical equation giving values of  $A$  in terms of well-known physical constants. Their complete equation may be written

$$\eta = 1 + \frac{\epsilon \sqrt{N} v_1 z_1}{30\eta_0 \sqrt{1000DkT} (z_1 + z_2) 4\pi} \times \left[ \frac{1}{2} \frac{\mu_1 z_2 + \mu_2 z_1}{\mu_1 \mu_2} - \frac{z_1 z_2 (\mu_1 - \mu_2)^2}{\mu_1 \mu_2 (\sqrt{\mu_1 z_1 + \mu_2 z_2} + \sqrt{(\mu_1 + \mu_2)(z_1 + z_2)})^2} \right] \sqrt{c},$$

where

$N$  = Avogadro's number

$v_1, v_2$  = numbers of ions

$z_1, z_2$  = valencies of ions

$\mu_1, \mu_2$  = absolute mobilities of ions

$D$  = dielectric constant of solvent

$k$  = Boltzmann's constant

$\epsilon$  = electronic charge

$\eta_0$  = viscosity of solvent

$T$  = absolute temperature.

\* 'J. Amer. Chem. Soc.,' vol. 51, p. 2950 (1929).

† 'Z. Phys. Chem.,' B, vol. 6, p. 159 (1929); 'Phys. Z.,' vol. 30, p. 611 (1929);  
'Phys. Z.,' vol. 32, p. 745 (1931); 'Phil. Mag.,' vol. 14, p. 537 (1932).

The validity of this equation has been established in aqueous solution for a series of uni-univalent electrolytes at two temperatures by Joy and Wolfenden\* and also to a higher degree of accuracy at one temperature by Jones and Talley.† The latter workers also showed that the coefficient *A* is zero for solutions of two non-electrolytes, sucrose and urea.

The object of the present work was to extend the experimental test of the equation to aqueous solutions of polyvalent electrolytes and to non-aqueous electrolyte solutions. The "B" coefficient in the Jones-Dole equation has also been studied and shown to lead to conclusions concerning the factors that determine its magnitude.

#### *Experimental Technique.*

In order to achieve the precision necessary for the measurement of the viscosity of very dilute solutions, a new modification of the capillary viscometer has been devised. By the simultaneous use in the same thermostat of twin viscometers of this type, one containing the solution and the other the pure solvent, differential measurements are possible in which important sources of error are eliminated by compensation.

In the modified form of viscometer illustrated in fig. 1 the U-tube is abolished. The pipette *A* and capillary *B* are mounted coaxially with the reservoir bulb *C* and held firmly in position by a large ground joint *D* at the top of the viscometer. The advantages of this design are

that the total time of flow has been increased, in comparison with the Washburn-Williams viscometer, and that although the capacity of the pipette is increased and the capillary lengthened no loss of rigidity is entailed.

The capillary tube is in two sections: the first is a short length, 1.5 mm. in diameter, which carries the lower reading line; the second and main section is narrower and is continued almost to the bottom of the reservoir. The

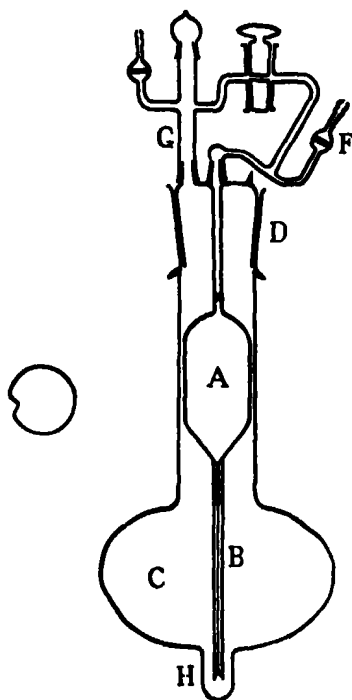


FIG. 1.—The Viscometer (inset: horizontal section of pipette *A*).

\* 'Proc. Roy. Soc.,' A, vol. 124, p. 413 (1931).

† 'J. Amer. Chem. Soc.,' vol. 55, p. 624 (1933).



upper structure of the viscometer, attached to D by two small ground joints, contains two sintered glass air-filters (E and F) and a broad tube G through which additions and abstractions of liquid may be made. One side of the pipette was slightly collapsed (*vide* horizontal section of pipette in fig. 1) to allow the canula of the Krogh pipette to reach the reservoir.

The dimensions and characteristics of the twin viscometers are summarized in Table I.

Table I.

Capacity of pipette bulb	100 c.c.
Capacity of reservoir	550 c.c.
Diameter of capillary	0.055 cm.
Length of capillary	12 cm.
Mean capillary velocity (water at 18° C.)	3.2 cm. per sec.
Time of flow (water at 18° C.)	4250 sec.
Change in time of flow with change in volume of liquid in viscometer	3.05 sec. per c.c.

The twin viscometers, immersed in a thermostat, were filled with pure solvent, the amounts being adjusted to give approximately equal times of flow. The solvent was then blown up into each pipette by pressure and the pressure equalized by opening the tap. The small time-difference between the passage of the two menisci past the upper reading lines was measured with a Venner stop-watch with 1/30 second escapement. The total time of flow was measured (for either viscometer) and finally the small time-difference between the passage of the two menisci past the lower reading-line. In these circumstances the ratio of the times of flow in the two viscometers is relatively insensitive to errors in the measurement of the total time of flow ; so long as the small time-differences, which never exceed 20 seconds, are measured accurately a much lower absolute precision is necessary in the measurement of the total time of flow.

A series of three or four such runs was necessary to determine the exact ratio of times of flow for the two viscometers containing pure solvent. With the aid of a Krogh\* syringe pipette, fitted with a long stainless steel canula, a small volume (*ca.* 0.5 c.c.) of a solution of known concentration was added to one of the two viscometers. This was completely mixed by a slow stream of bubbles of solvent-saturated air from the capillary and then the Krogh pipette was used to abstract an exactly equal volume of the diluted solution. Using the Krogh pipette additions and abstractions may be made with a

\* Krogh and Keys, 'J. Chem. Soc.,' vol. 133, p. 2436 (1931).

precision of 0.0005 c.c. so that an exact amount of concentrated solution may be added without any change in the hydrostatic head of the viscometer.

The ratio of times of flow for the two viscometers was now measured again in a series of three or four runs. A second addition of concentrated solution was now made and the process repeated. As many as four or five additions were made to one viscometer while the other, still containing pure solvent, was used for comparison. When a series of concentrations had been measured in the first viscometer, this viscometer was used as a standard and a second series of additions of concentrated solution was made to the pure solvent in the second viscometer.

By the differential use of twin viscometers errors due to variation of temperature and barometric pressure are eliminated. The uncertainty introduced by timing errors is unlikely to exceed one part in forty thousand. The difference between the surface tension of the solution and that of the pure solvent is negligible for solutions more dilute than N/50; the kinetic energy correction never exceeded one part in one hundred thousand. The error which might arise owing to the minute change in hydrostatic head caused by the contraction of a concentrated solution on dilution did not exceed the timing error at the concentrations measured.

The salts used were well-crystallized pure specimens; the lanthanum chloride was shown to contain less than 0.3% of cerium. Nitrobenzene was prepared by the nitration of carefully purified benzene, dried with calcium chloride and phosphorus pentoxide and finally distilled in an all-glass still under reduced pressure at 75° C. Ethyl alcohol was kindly supplied to us by Mr. O. L. Hughes of this laboratory; it had been refluxed over lime and then distilled successively from aluminium ethoxide and anhydrous copper sulphate. When non-aqueous solvents were being used in the viscometers, they were protected from atmospheric moisture by drying-tubes containing phosphorus pentoxide and attached directly to the viscometers by vapourless cement. With these precautions the time of flow in the viscometers remained steady over a period of 10 days.

### *Experimental Results.*

The relative viscosity of the solutions was calculated in the following way. Let the mean ratio of the times of flow of the pure solvent in both viscometers be  $\left( \frac{t \text{ solvent I}}{t \text{ solvent II}} \right)$  and let the mean ratio of the times of flow with the first

viscometer containing a solution be  $\left(\frac{t \text{ solution I}}{t \text{ solvent II}}\right)$ . Correcting for the difference between the apparent density in air of the solution and that of the pure solvent, the relative viscosity of the solution is given by the expression :

$$\eta = \left(\frac{t \text{ solvent II}}{t \text{ solvent I}}\right) \left(\frac{t \text{ solution I}}{t \text{ solvent II}}\right) \times \begin{array}{l} \text{apparent density of solution in air relative} \\ \text{to solvent at the same temperature.} \end{array}$$

Except for magnesium sulphate, where the data of Lamb and Lee\* were available, density measurements were made by the technique described by Hartley and Barrett.† The viscosity of each solution was calculated from densities graphically interpolated from the results in Table II.

Table II.

Salt.	Solvent.	Concentration (in moles per litre).	Apparent density of solution in air.
Lanthanum chloride (at 18° C.)	Water	0.0333 0.00067	1.00755 1.00151
Tetraethylammonium picrate (at 18° C.)	Water	0.0183 0.00918	1.00179 1.00089
	Nitrobenzene	0.0500 0.0250	1.00141 1.00067
Sodium iodide (at 25° C.)	Ethyl alcohol	0.1061 0.0265	1.01818 1.00455

In Table III the concentration of the solution, expressed in moles per litre, is given in column 1, the relative viscosity of the solution, the mean of two or more independent determinations in column 2, and the function  $(\eta - 1)/\sqrt{c}$  is tabulated in column 3.

### Discussion of Results.

By plotting the function  $(\eta - 1)/\sqrt{c}$  against  $\sqrt{c}$  the results embodied in the above tables may be used to evaluate the two coefficients of the Jones-Dole equation for the electrolytes studied. The intercept on the axis of zero concentration is equal to A and the slope of the line is equal to B. The results are plotted in this way in fig. 2; the circle round each experimental point represents the probable error.

\* 'J. Amer. Chem. Soc.,' vol. 35, p. 1666 (1913).

† 'J. Chem. Soc.,' vol. 99, p. 1072 (1911).

Table III.

Concentration (in moles per litre).	Mean relative viscosity.	$\frac{\eta - 1}{\sqrt{c}}$
<b>I. Magnesium sulphate in water at 18° C.</b>		
0.000325	1.00038 <sub>6</sub>	0.021 <sub>7</sub>
0.000735	1.00076 <sub>6</sub>	0.028 <sub>6</sub>
0.000995	1.00101	0.032 <sub>6</sub>
0.00138	1.00134	0.036 <sub>6</sub>
0.00197	1.00189	0.042 <sub>6</sub>
0.00377	1.00324	0.052 <sub>7</sub>
0.00434	1.00385	0.058 <sub>6</sub>
0.00610	1.00499	0.063 <sub>6</sub>
0.00905	1.00720	0.075 <sub>7</sub>
0.01310	1.01001	0.087 <sub>6</sub>
<b>II. Lanthanum chloride in water at 18° C.</b>		
0.000217	1.00067	0.045 <sub>6</sub>
0.000254	1.00070	0.043 <sub>6</sub>
0.00065	1.00130	0.050 <sub>6</sub>
0.00094	1.00162	0.052 <sub>6</sub>
0.00130	1.00206	0.057 <sub>6</sub>
0.00211	1.00276	0.060 <sub>6</sub>
0.00289	1.00351	0.065 <sub>6</sub>
0.00437	1.00484	0.073 <sub>6</sub>
0.00607	1.00579	0.074 <sub>6</sub>
<b>III. Tetraethylammonium picrate in water at 18° C.</b>		
0.000622	1.00071	0.028 <sub>6</sub>
0.000724	1.00084	0.031 <sub>6</sub>
0.00122	1.00127	0.036 <sub>6</sub>
0.00140	1.00140	0.037 <sub>6</sub>
0.00207	1.00204	0.044 <sub>6</sub>
0.00236	1.00224	0.046 <sub>6</sub>
0.00272	1.00263	0.050 <sub>6</sub>
0.00393	1.00352	0.056 <sub>6</sub>
0.00504	1.00452	0.063 <sub>6</sub>
<b>IV. Tetraethylammonium picrate in nitrobenzene at 18° C.</b>		
0.000537	1.00091	0.039 <sub>6</sub>
0.000889	1.00127	0.042 <sub>6</sub>
0.00107	1.00152	0.046 <sub>6</sub>
0.00213	1.00290	0.062 <sub>6</sub>
0.00265	1.00342	0.066 <sub>6</sub>
0.00425	1.00529	0.081 <sub>6</sub>
0.00441	1.00541	0.081 <sub>6</sub>
0.00633	1.00759	0.095 <sub>6</sub>
0.00790	1.00935	0.105 <sub>6</sub>
<b>V. Sodium iodide in ethyl alcohol at 25° C.</b>		
0.00070	1.00146	0.055 <sub>6</sub>
0.00087	1.00183	0.062 <sub>6</sub>
0.00100	1.00207	0.065 <sub>6</sub>
0.00139	1.00265	0.071 <sub>6</sub>
0.00301	1.00485	0.088 <sub>6</sub>
0.00347	1.00566	0.096 <sub>6</sub>
0.00487	1.00739	0.105 <sub>6</sub>
0.00601	1.00879	0.113 <sub>6</sub>
0.00691	1.00993	0.119 <sub>6</sub>
0.00900	1.01220	0.128 <sub>6</sub>
0.01396	1.01755	0.148 <sub>6</sub>
0.01680	1.02079	0.160 <sub>6</sub>

The graphs for tetraethylammonium picrate and for lanthanum chloride are linear over the whole concentration range. The graph for magnesium sulphate begins to show curvature somewhere above N/200 while that for sodium iodide is unequivocally straight over no part of the concentration range measured. Values of A and B have been determined graphically and confirmed where

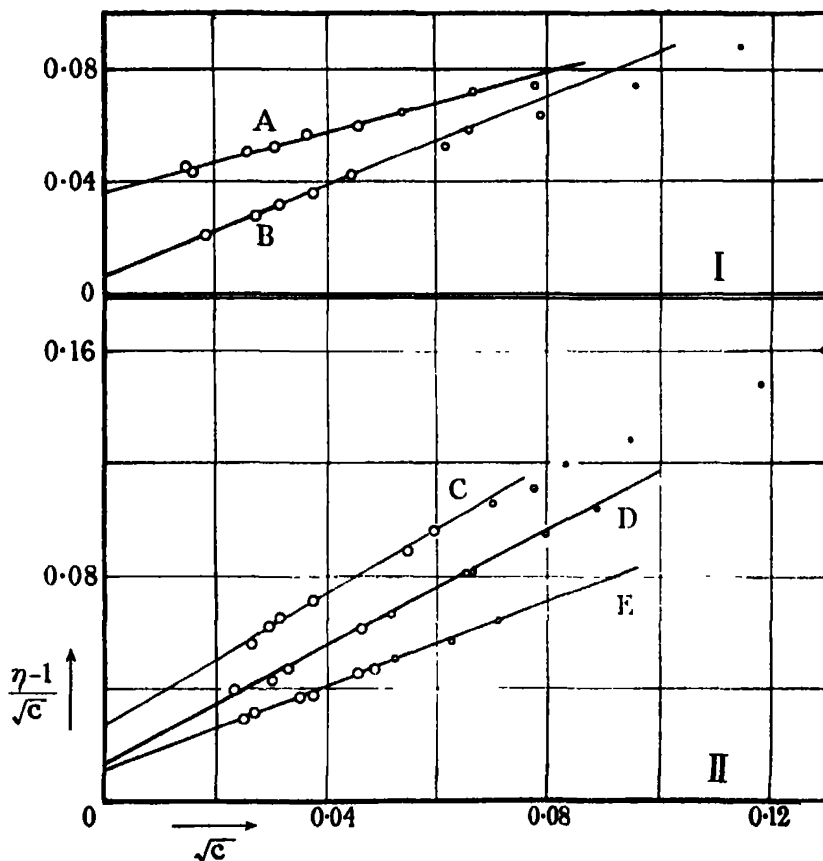


FIG. 2.—I: Upper set of co-ordinates—A, lanthanum chloride in water at 18° C.; B, magnesium sulphate in water at 18° C. II: Lower set of co-ordinates—C, sodium iodide in ethyl alcohol at 25° C.; D, tetraethylammonium picrate in nitrobenzene at 18° C.; E, tetraethylammonium picrate in water at 18° C.

possible by the method of least squares; the results are embodied in Table IV.

Although the viscosity measurements themselves are more accurate than those of Joy and Wolfenden, the high values of the coefficient B for the electrolytes measured in the present work make the plots of  $(\eta - 1)/\sqrt{c}$  so much steeper in slope that the extrapolated values of the intercept A are no more

accurate than the values they were able to obtain. The experimental results for lanthanum chloride were never so consistent as those obtained with other electrolytes; it is possible that the anomalies were due to traces of hydrolysis although the addition of hydrochloric acid to repress hydrolysis produced no appreciable change in viscosity.

Table IV.

Salt.	Solvent.	Temperature °C.	A Experimental.	A Theoretical.	B.
Magnesium sulphate	Water	18	$0.0066 \pm 0.0005$	0.0023	0.808
Lanthanum chloride	Water	18	$0.037 \pm 0.002$	0.0318	0.516
Tetraethyl- ammonium picrate	Water	18	$0.0108 \pm 0.002$	0.0118	0.748
	Nitrobenzene	18	$0.0129 \pm 0.002$	0.0166*	1.039
Sodium iodide	Ethyl alcohol	25	$0.027 \pm 0.008$	0.0255	1.15

\* The conductivity of tetraethylammonium picrate in nitrobenzene does not agree with Walden's rule. If mobilities calculated from this rule instead of from the conductivity data are used, a value of 0.0154 is obtained for the coefficient A.

*The Coefficient A.*—Aqueous solutions of magnesium sulphate and lanthanum chloride do not give results in accordance with the Falkenhagen equation and these electrolytes are the first cases to be investigated which have shown an unmistakable divergence. Such a divergence is not unexpected as it is characteristic of the behaviour of electrolytes of higher valency-type in respect to other properties such as those of heats of dilution, etc. The discrepancy is presumably associated with the "higher terms" of the ionic atmosphere theory.

In water tetraethylammonium picrate shows excellent agreement with the theory. In nitrobenzene the agreement is not so good but the experimental value of A is greater than in aqueous solution as the theory demands. It is more difficult to compare theory and experiment for the alcoholic solutions of sodium iodide since the line from which the extrapolation is made seems to show slight curvature down to the lowest concentrations measured; the experimental value of A is, however, consistent with the theoretical value. The measurements in nitrobenzene and ethyl alcohol cannot be regarded as a conclusive test of the dielectric constant term in the Falkenhagen equation but they show that the A coefficient in both solvents is of the same order as the theoretical value.

*The Coefficient B.*—The coefficient B in the Jones-Dole equation has heretofore received little attention. Whereas the coefficient A can only be satisfactorily evaluated from measurements at high dilutions, the coefficient B can be found to a fair degree of accuracy from viscosity data obtained at higher concentrations. For uni-univalent salts which raise the viscosity of

Table V.—Values of B at 18° C.

	IO <sub>3</sub> .	Cl.	I.	OH.	NO <sub>3</sub> .	SO <sub>4</sub> .	ClO <sub>3</sub> .	Picrate.
Li .. ..	0.28 <sub>3</sub>	0.13 <sub>3</sub>	—	—	0.08 <sub>3</sub>	0.25	—	—
Na .. ..	0.23	0.07 <sub>4</sub>	—	0.19 <sub>3</sub>	0.03 <sub>1</sub>	—	—	—
K .. ..	0.11	-0.03 <sub>3</sub>	-0.10 <sub>1</sub>	0.09 <sub>1</sub>	-0.07	0.08	-0.05 <sub>7</sub>	—
Rb .. ..	—	-0.00 <sub>3</sub>	—	—	-0.11 <sub>1</sub>	—	—	—
Cs .. ..	—	-0.07 <sub>3</sub>	—	—	-0.12 <sub>2</sub>	—	—	—
NH <sub>4</sub> ..	—	-0.05	—	—	—	—	—	—
H .. ..	—	—	—	—	0.00 <sub>3</sub>	—	—	—
N(Et) <sub>4</sub>	—	—	—	—	—	—	—	0.74 <sub>2</sub>
Mg .. ..	—	—	—	—	—	0.80 <sub>3</sub>	—	—
La .. ..	—	0.51 <sub>4</sub>	—	—	—	—	—	—

water the Jones-Dole equation is usually obeyed satisfactorily up to a concentration of about  $N/2$  and sometimes further; values of B can therefore be found directly in such cases from existing data. For salts which lower the viscosity of water the plot of  $(\eta - 1)/\sqrt{c}$  against  $\sqrt{c}$  usually shows curvature above  $N/10$ ; however, since the intercept on the axis of zero concentration

Table VI.—Values of B at 25° C.

	Cl.	Br.	I.	ClO <sub>3</sub> .	BrO <sub>3</sub> .	NO <sub>3</sub> .	SO <sub>4</sub> .
Li .. ..	—	—	—	—	—	0.10 <sub>1</sub>	—
K .. ..	-0.01 <sub>4</sub>	—	—	-0.03 <sub>1</sub>	-0.00 <sub>1</sub>	-0.05 <sub>1</sub>	—
Rb .. ..	-0.03 <sub>7</sub>	-0.06 <sub>1</sub>	-0.11	—	—	—	—
Cs .. ..	-0.05 <sub>0</sub>	—	—	—	—	-0.09 <sub>3</sub>	—
NH <sub>4</sub> ..	-0.14 <sub>3</sub>	-0.03 <sub>7</sub>	-0.08	—	—	-0.03 <sub>4</sub>	0.20
N(Et) <sub>4</sub>	—	—	—	—	—	—	—
H .. ..	0.06	0.02 <sub>3</sub>	—	0.04 <sub>3</sub>	—	—	0.33

(= A) may be calculated from the Falkenhagen equation, it is possible to estimate graphically the limiting slope at high dilutions from experimental work at higher concentrations.

By the application of these methods to the more reliable sources of viscosity data, a provisional series of values of B has been calculated at 18° C. and 25° C. (Table V and VI). The data have been taken from the results of the

following workers: Grüneisen,\* Applebey,† Merton,‡ Joy and Wolfenden, Jones and Talley, and the present work, except for a few substances for which data have been taken from the International Critical Tables.

Since the coefficient A accounts for the interionic forces, the coefficient B is presumably due to other effects characteristic of the individual ions of the electrolyte and might be expected to be made up additively of two independent coefficients characteristic of anion and cation respectively. An examination of the values of B for a series of salts of two univalent metals shows that there is a constant difference between the value of B for the salts of two metals with a common anion. This result is illustrated in Table VII.

Table VII.

	IO <sub>3</sub> .	Cl.	NO <sub>3</sub> .	OH.	SO <sub>4</sub> /2.
Na — K	0.12 ± 0.015	0.107 ± 0.004	0.107 ± 0.003	0.105 ± 0.01	—
K — Rb	—	0.032 ± 0.005	0.037 ± 0.01	—	—
Rb — Cs	—	0.010 ± 0.005	0.015 ± 0.005	—	—
Li — K	0.17 ± 0.015	0.172 ± 0.005	0.158 ± 0.01	—	0.171 ± 0.01

This table suggests that for uni-univalent electrolytes B is an additive quantity within the experimental error. Further experimental work at present in progress will be necessary before the additivity can be regarded as more than a first approximation. In what follows, by making the provisional assumption of additivity and considering the B coefficients of individual ions it is possible to show parallelisms which throw some light on the nature of B. These parallelisms would be much less clear if the joint (and sometimes antagonistic) effects of anion and cation were taken together but the inferences in the latter part of this paper are not necessarily dependent on the strict additivity of B.

In order to divide B into components for the individual ions some arbitrary assumption must be made about the nature of the B term. One possibility, which admittedly must make the picture too simple, is that large ions with their correspondingly weak fields of force behave like the particles of a suspensoid. Einstein§ has shown that the relative viscosity of a medium containing spherical particles in suspension is given by the equation

$$\eta = 1 + 2.5v,$$

\* 'Wiss. Abh. phys.-tech. Reichsanst. Berl.,' vol. 4, p. 237 (1905).

† 'J. Chem. Soc.,' vol. 97, p. 2000 (1910).

‡ 'J. Chem. Soc.,' vol. 97, p. 2454 (1910).

§ 'Ann. Physik.,' vol. 19, p. 289 (1906); vol. 34, p. 591 (1911).



where  $\eta$  is the relative viscosity of the suspensoid, and  $v$  is the aggregate volume of the particles contained in 1 c.c. of the medium. This equation postulates that the spheres are large compared with the molecules of the suspending medium so that the parallelism between  $B$  and the linear term of the Einstein equation only assumes some degree of plausibility for ions of considerable size. If, however, we restrict ourselves to ions whose temperature coefficient of mobility corresponds to the variation of fluidity of the solvent in the manner prescribed by Stokes's law (*i.e.*, to ions of which the mobility, under a potential gradient, is consistent with Stokes's picture of a sphere moving through a continuous medium) then there is some justification for the expedient of assuming  $B$ , like the term in the Einstein equation, to be proportional to the ionic volume.

The variation of the mobility of the lithium and the iodate ions with temperature is in conformity with Stokes's law (*vide* p. 487), and therefore we assume that the individual  $B$  coefficients of the ions in lithium iodate may be regarded as proportional to the ionic volumes. Taking the ionic volumes as inversely proportional to the cube of the ionic mobilities, we find (since  $u_{Li} = 32.75$  and  $u_{IO_3} = 33.90$  at  $18^\circ \text{C}$ .) that  $B_{Li} = 0.146$  and  $B_{IO_3} = 0.136$ . Starting from these values Table VIII of individual  $B$  coefficients may be drawn up.

Table VIII.

Ion.	$B_{18^\circ \text{C}}$	Ion.	$B_{18^\circ \text{C}}$
Li	0.14 <sub>6</sub>	OH	0.11 <sub>2</sub>
Na ..	0.08 <sub>4</sub>	Cl	-0.00 <sub>7</sub>
K	-0.02 <sub>5</sub>	NO <sub>3</sub>	-0.06 <sub>1</sub>
Rb	-0.05 <sub>6</sub>	*Picrate	0.43 <sub>4</sub>
Ca	-0.06 <sub>6</sub>	ClO <sub>4</sub>	-0.03 <sub>1</sub>
*N(Et) <sub>4</sub>	0.31 <sub>4</sub>	IO <sub>3</sub>	0.13 <sub>6</sub>
H	0.06 <sub>6</sub>		

\* It is interesting to compare these values with those of the  $B$  coefficient of sucrose at  $25^\circ \text{C}$ . evaluated by Jones and Talley as 0.8786.

Another series of values can be calculated at  $25^\circ \text{C}$ ., if it is assumed that  $B_{Li}$  is unchanged by the small rise in temperature. This assumption is justified by the fact that the temperature coefficient of mobility of the lithium ion is in conformity with Stokes's law. Again taking  $B_{Li} = 0.146$  we obtain the values given in Table IX.

The parallelism between the  $B$  coefficient and the corresponding term in the Einstein equation is only plausible for large ions. For other ions complicating factors must be taken into account; the existence of negative values of  $B$

indicates how large a part these factors may play. The generally accepted explanation of "negative viscosity," corresponding to negative values of  $B$ , is the depolymerization\* of the solvent by the ions, and a parallel is drawn between the diminished viscosity of water under high pressures and that due to the electrostrictive forces of the ions in a solution. This view is supported by a consideration of the temperature coefficient of  $B$  for individual ions.

Table IX.

Ion.	$B_{25^\circ \text{C.}}$	Ion.	$B_{25^\circ \text{C.}}$
Li	0.14 <sub>8</sub>	NO <sub>3</sub>	-0.04 <sub>8</sub>
K	-0.00 <sub>8</sub>	ClO <sub>3</sub>	-0.02 <sub>8</sub>
Rb	-0.03 <sub>8</sub>	BrO <sub>3</sub>	0.00 <sub>7</sub>
Cs	-0.04 <sub>8</sub>	Cl	-0.00 <sub>8</sub>
H	0.06 <sub>9</sub>	Br	-0.03 <sub>1</sub>
NH <sub>4</sub>	-0.00 <sub>8</sub>	I	-0.07 <sub>9</sub>

At low temperatures water is considerably polymerized so that it is possible for ions to exert a very marked depolymerizing action. At higher temperatures water is already thermally depolymerized to a large extent and therefore the possible depolymerizing effect of the ions is correspondingly reduced. There is therefore a tendency for negative  $B$  coefficients to increase algebraically with rise of temperature, as is shown in Table X.

Table X.

Ion.	$B_{15^\circ \text{C.}}$	$B_{25} - B_{15}$
Li	0.14 <sub>8</sub>	0.000
H	0.06 <sub>9</sub>	0.000
Cl	-0.00 <sub>7</sub>	0.00 <sub>1</sub>
K	-0.02 <sub>8</sub>	0.01 <sub>7</sub>
ClO <sub>3</sub>	-0.03 <sub>1</sub>	0.00 <sub>8</sub>
Rb	-0.05 <sub>8</sub>	0.02 <sub>8</sub>
NO <sub>3</sub>	-0.06 <sub>1</sub>	0.01 <sub>8</sub>
Cs	-0.06 <sub>8</sub>	0.02 <sub>9</sub>

If the solvent in the immediate neighbourhood of an ion is partially depolymerized by electrostriction, the temperature coefficient of viscosity of the solvent in the neighbourhood of the ion will be lower than that of the solvent

\* In terms of the hypothesis, recently put forward by Bernal and Fowler ('J. Chem. Phys.,' vol. 1, p. 515 (1933)), of water as a "pseudo-crystalline" liquid, the word "depolymerization" may be replaced by "rise of structural temperature."

as a whole. Since the mobility of an ion depends on the viscosity of the solvent immediately surrounding it, the temperature coefficient of mobility of an ion must be determined by the (lower) temperature coefficient of viscosity of the partially depolymerized solvent and not by the temperature coefficient of macroscopic viscosity. This provides a clue to the well-known anomaly\* that with increase of temperature the mobility of most ions increases less rapidly than the viscosity of the solvent diminishes. This anomaly, which corresponds in terms of Stokes's law to an increase in the apparent ionic radius with rise of temperature, is immediately accounted for by the lowered temperature coefficient of viscosity in the neighbourhood of the ion.

If this explanation is true it is to be expected that the anomaly coefficient  $F$  defined by the equation

$$F_{T_1} = \frac{\eta_{T_1} \times \mu_{T_1}}{\eta_{T_2} \times \mu_{T_2}},$$

where

$\eta$  = viscosity of solvent

$\mu$  = mobility of ion

$T_1$  and  $T_2$  are two temperatures ( $T_1 < T_2$ ),

would show the highest values for ions whose  $B$  coefficients are negative. Such a correlation is illustrated by Table XI based on the results of Walden and Ulich.†

Table XI.

Ion.	$F_{10}^{100}$ .	$B_{10}$ .
Li	0.997	0.146
Na	1.049	0.084
K	1.198	-0.025
Cs	1.254	-0.065
N(Et) <sub>4</sub>	1.002	0.314
Cl	1.160	-0.007
NO <sub>2</sub>	1.221	-0.061
Pierate	1.000	0.434
ClO <sub>3</sub>	1.193	-0.031
IO <sub>3</sub>	1.008	0.136

The idea that the mobility of an ion is determined by a "microscopic" viscosity of the solvent in its immediate neighbourhood also offers an explanation of the fact that the radii of certain ions in water (where they are certainly

\* This anomaly is discussed very thoroughly by Ulich ('Fortschr. Chem. Phys.,' vol. 18, p. 14 (1926) and ('Hand-und Jahrb., Chem. Physik,' vol. 6, pp. 143 and 181 (1933)).

† 'Z. phys. Chem.,' vol. 107, p. 219 (1923).

hydrated), when calculated from their mobilities and the *macroscopic* viscosity of the solvent by means of Stokes's law, appear to be smaller than their lattice radii.

Table XII.

Ion.	$R_{18}^{\circ}$ .	Lattice radius.	"Stokes' radius" in water at 18°.
		A.	A.
Li ...	0.146	0.72	2.30
Na ...	0.084	1.01	1.79
K ...	-0.025	1.30	1.22
Cs ...	-0.066	1.68	1.11
Cl ...	-0.007	1.72	1.21
Br ...	-0.031	(at 25°) 1.92	1.18
I ...	-0.079	(at 25°) 2.19	1.20
N(Et) <sub>4</sub>	0.314	—	2.67

The authors are indebted to Imperial Chemical Industries and to the Chemical Society for grants which defrayed the cost of part of the apparatus.

### *Summary.*

(1) A modified capillary viscometer is described which, when used in conjunction with a new differential technique, may be employed to measure the relative viscosity of dilute solutions to one part in forty thousand.

(2) The viscosity of aqueous solutions of magnesium sulphate and lanthanum chloride, of tetraethylammonium picrate in water, and nitrobenzene, and of sodium iodide in ethyl alcohol have been measured over concentrations ranging from M/5000 to M/100. The results have been compared with the Falkenhagen equation.

(3) The coefficient B of the Jones-Dole equation has been evaluated for a number of uni-univalent electrolytes in aqueous solution, and has been shown, to a first approximation, to be made up additively of two quantities characteristic of anion and cation.

(4) The phenomenon of "negative viscosity" is discussed, and an explanation of the abnormal temperature coefficient of mobility in ions is suggested. It is also shown that Stokes's law must lead to erroneous results for a large number of ions.

## *The Constitution of Water in Different States.*

By I. RAMAKRISHNA RAO, M.A., Ph.D., Andhra University, Waltair.

(Communicated by O. W. Richardson, F.R.S.—Received October 18, 1933.)

### 1. *Introduction.*

In a previous communication,\* the author described the behaviour of water with changes of temperature and with addition of electrolytes as revealed by the changes in the structure of its Raman band. This band is found to consist of three components with mean Raman frequencies 3217, 3433, and 3582  $\text{cm.}^{-1}$ . With increasing temperature, the component 3582 increased in intensity, while the first diminished, the central one remaining nearly constant. A comparison of the above changes with the results obtained by the author for ice† indicated that the component which becomes more intense with increasing temperature is absent in the Raman band of ice. The first component with the smallest Raman frequency, however, is very conspicuous in ice, and the central one retains the same relative intensity.

The above differences in the structure of the band were explained on the hypothesis that water consists not only of single ( $\text{H}_2\text{O}$ ) but also of double ( $(\text{H}_2\text{O})_2$ ) and triple ( $(\text{H}_2\text{O})_3$ ) molecules, the relative proportions of which depend upon its state of aggregation as well as on its temperature. Ice is supposed to consist only of double and triple molecules, which reveal themselves as the two components of its Raman band. In the liquid state, the presence of all the three components in this band is taken as evidence of the fact that there are all the three types of molecules in this state. No data were then available as regards the Raman band for water-vapour. But the infra-red absorption band observed by Hettner‡ at 2.66  $\mu$  has its corresponding component in the water band which is supposed to arise out of the single  $\text{H}_2\text{O}$  molecules.

The above hypothesis was based on the results obtained with water only at two different temperatures 14° and 75° C., and with ice, of which the band was taken with a spectrograph of very small dispersion which could not resolve the components. Also the Raman band for water was unresolved, though the changes described above were very conspicuous. It was therefore thought

\* 'Proc. Roy. Soc.,' A, vol. 130, p. 489 (1931).

† 'Ind. J. Phys.,' vol. 3, p. 131 (1928).

‡ 'Ann. Physik,' vol. 55, p. 476 (1918).

desirable to extend the work over a wider range of temperatures and in all the three states. The results thus obtained are described below.

## 2. *Experimental Details.*

To obtain the Raman spectrum of ice, long rectangular blocks were kept in a double-walled wooden box open at the top packed between the walls with sawdust. Light from the mercury arc was condensed on to the block of ice by means of a cylindrical condenser of pyrex glass tubing 5 cm. in diameter, containing water. The light scattered by the whole length of the block was concentrated on to the slit of the spectrograph by a lens in the usual manner. With this, an exposure of about 10 hours gave a picture which was just good enough for microphotometering. Owing to the rapid melting of the ice, the blocks had to be replaced once every half an hour.

For experiments with water at different temperatures, the arrangement shown in fig. 1 was adopted. The Wood's tube W.T. was made of quartz

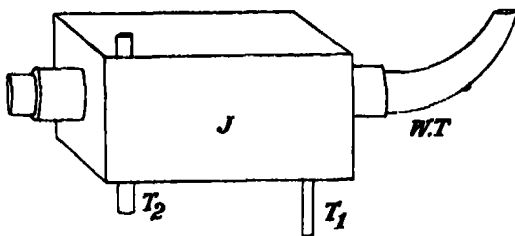


FIG. 1.

with a plane quartz window melted on to one end and bent into a horn at the other. The cooling jacket *J* consisted of a rectangular copper box open at the top. Two brass tubes, 3 cm. in length and diameter slightly larger than that of the Wood's tube, were soldered coaxially at the two square ends of the box, so that the quartz tube containing the liquid under investigation could be passed through them and fixed by means of wide rubber tubing. The cooling water entered the box through a narrow tube  $T_1$  soldered at the bottom, and after filling the box nearly up to the top, drained through a wider tube  $T_2$  soldered through the bottom of the box with its upper end nearly at the top. This end was about 3 mm. higher than the top of the quartz tube so that there was a layer of water about 3 mm. in thickness above it.

The mercury arc lamp was placed slightly above the box with its axis parallel to the quartz tube, thus illuminating the liquid directly. Two ellipsoidal metallic reflectors one above the lamp and the other below the quartz tube in

the box served to concentrate most of the light into the liquid. The other arrangements were as usual.

For working with water at 0° C., the water circulation was stopped. The two brass tubes were packed in cotton leaving only the quartz window open. Ice was constantly added to the water contained in the jacket, which was kept stirred. A thermometer, introduced into the quartz tube from the side of the horn, indicated the temperature of the water contained in it. When it attained 0° C. the exposure was started, and the temperature maintained at 0° C. by continued addition of broken ice. A similar procedure was adopted for maintaining the water at 4° C.

To obtain the Raman spectrum of water at the room temperature, the water circulation was maintained. For water at 100° C., however, a different method had to be used. Nichrome wire was wound round the quartz tube which was taken out of the copper jacket. The windings were at a distance from one another so that most of the tube was visible and the water inside could be illuminated by the mercury lamp placed directly above it. The bottom of the quartz tube was surrounded by asbestos sheet to prevent cooling. The heat from the mercury arc itself served to eliminate cooling from above. The current through the nichrome wire was so adjusted that the temperature of the water, under steady conditions, was about 2° C. below the boiling point. With such an arrangement the temperature was constant to within  $98^{\circ}\text{C.} \pm 1^{\circ}$ .

The author intended to study the Raman spectrum of water-vapour also, but as Daure and Kastler\* have published their results with this substance, the value for the Raman frequency obtained by them is assumed in this investigation.

The method of determining intensities, which is adopted in this communication, was described in detail in another paper by the author.†

### 3. Results.

The micro-photometric curves of the Raman spectra of ice and water at different temperatures are reproduced in fig. 2. The wave-lengths of the mercury arc lines are marked and the Raman bands of water indicated by arrows. The left one is excited by the 3650-63 group and the right by the 4047 line. The curve for ice clearly shows that each of the bands consists of two components the one towards the left being sharper and also greater in intensity. Water at

\* 'C. R. Acad. Sci. Paris,' vol. 192, p. 1721 (1931).

† 'Proc. R. Acad., Sci., Amst.,' vol. 33, p. 638 (1930).

0° C. also gives rise to two components which are not as clearly resolved from one another as in ice. There is also a change in the position of the components, which have slightly shifted to the right, and the one that is sharp in ice has become diffuse on changing to water. The sharp band of greater intensity in ice has become, in water at 0° C., nearly equal in intensity to the other component.

There is no perceptible difference between the water bands at 0° C. and 4° C., except for a slight further diminution in the intensity of the component to the left, their position being nearly the same at both temperatures. At 38°

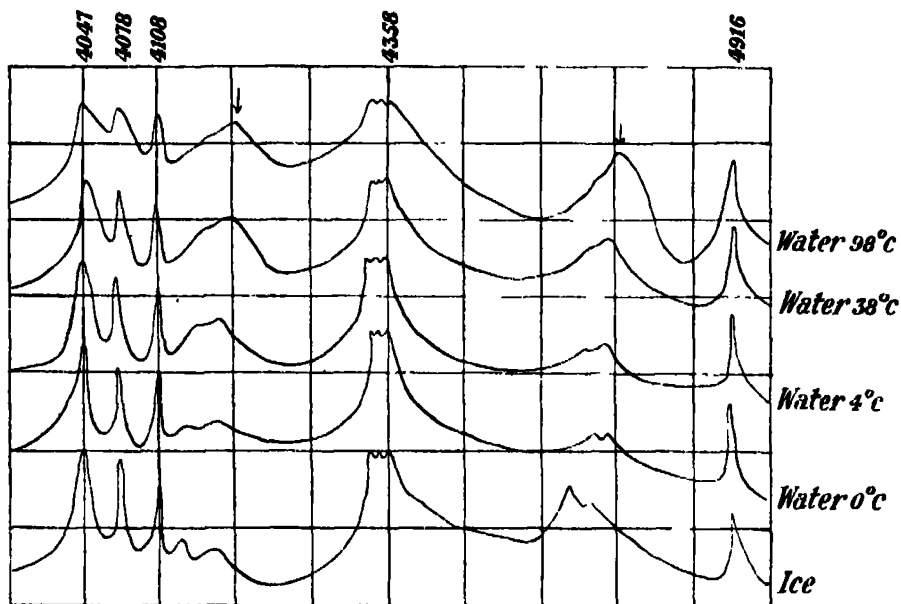


FIG. 2.—Microphotometer curves of Raman bands of ice and water at different temperatures.

however, there is clearly a change in the form of the microphotometric curve corresponding to the band. Of the two maxima which are clear in the curves for water at 0° C. and 4° C., the left component has considerably diminished in intensity. The topmost curve corresponding to water at 98° C. is very much different in appearance from the others. There is a large maximum whose position is shifted towards the longer wave-length side compared to the maxima of the other curves.

The following table contains the Raman frequencies of  $H_2O$  calculated from the microphotometric curves given above. As the bands are very diffuse, the exact position of the maxima could not be located for measurement with a micrometer. Hence the wave-lengths are determined by measuring distances on



the photometric curves themselves using the mercury arc lines as the standards. As such, the values given are not very accurate. Since, however, it is the intensities that are more important in this work, a high degree of accuracy in measurement of wave-lengths is not found to be necessary. The values given in the table are for the band excited by the 4047 line alone. The other due to excitation by the 3650-63 group is a mixture of the bands arising from the three components of the group and hence is not simple enough to give accurate values of the frequency.

Table I.—Raman Frequencies of the  $\text{H}_2\text{O}$  Band in different States.

State.	Raman frequencies of maxima in $\text{cm}^{-1}$ .			
Ice	—	—	3321 (3.01 $\mu$ )	3196 (3.13 $\mu$ )
Water 0° C.	—	3502 (2.86 $\mu$ )	3321 (3.01 $\mu$ )	—
Water 4° C.	—	3412 (2.93 $\mu$ )	3284 (3.06 $\mu$ )	—
Water 38° C.	—	3493 (2.86 $\mu$ )	3311 (3.02 $\mu$ )	—
Water 98° C.	—	3466 (2.89 $\mu$ )	—	—
Water-vapour	3655 (2.74 $\mu$ )	—	—	—

The first column in Table I contains the state of  $\text{H}_2\text{O}$ , the other columns the Raman frequencies of the maxima of the bands. The values given within brackets below each value of  $\delta\nu$  are the corresponding infra-red wave-lengths in  $\mu$ .

As it is not possible to draw any conclusions from the simple microphotometric curves, the intensities of the bands are calculated according to the method already described. The following Tables II-VI contain the calculated results of the intensity distribution of the band excited by the 4047 line in each of the above states.

The results given in Tables II-VI are represented graphically in fig. 3. The wave-lengths are taken along the abscissa and the intensities along the ordinate. The features of the intensity distribution of the bands described from the microphotometric curves are much more pronounced in the intensity curves. The solid line corresponds to ice and consists, as already mentioned, of two distinct maxima at 3.01 and 3.13  $\mu$ , the latter being much stronger in intensity

Table II.—Intensity Distribution of the H<sub>2</sub>O Band of Ice.

$\lambda$ in A.U.	$\nu$ in cm. <sup>-1</sup> .	$\delta\nu$ in cm. <sup>-1</sup> .	$\lambda$ I.R. in $\mu$ .	Intensity.
4775	20937	3768	2.65	0
4745	21069	3636	2.75	5
4723	21167	3538	2.83	12
4707	21239	3466	2.89	30
4690	21316	3389	2.95	39
4675	21384	3321	3.01	40
4660	21453	3252	3.07	37
4648	21509	3196	3.13	55
4632	21583	3122	3.20	36
4610	21666	3019	3.31	10
4580	21828	2877	3.48	0

Table III.—Intensity Distribution of H<sub>2</sub>O Band. Water at 0° C.

$\lambda$ in A.U.	$\nu$ in cm. <sup>-1</sup> .	$\delta\nu$ in cm. <sup>-1</sup> .	$\lambda$ I.R. in $\mu$ .	Intensity.
4778	20923	3782	2.64	6
4750	21020	3685	2.71	11
4732	21127	3578	2.80	13
4715	21203	3502	2.86	48
4693	21302	3403	2.94	41
4675	21384	3321	3.01	47
4654	21481	3224	3.10	30
4636	21564	3141	3.18	18
4614	21667	3038	3.29	6

Table IV.—Intensity Distribution of the H<sub>2</sub>O Band of Water at 4° C.

$\lambda$ in A.U.	$\nu$ in cm. <sup>-1</sup> .	$\delta\nu$ in cm. <sup>-1</sup> .	$\lambda$ I.R. in $\mu$ .	Intensity.
4771	20954	3751	2.67	9
4752	21042	3663	2.73	18
4730	21136	3569	2.80	35
4708	21235	3470	2.88	44
4695	21293	3412	2.93	49
4680	21362	3343	2.99	36
4667	21421	3284	3.05	38
4652	21490	3215	3.11	31
4635	21569	3136	3.19	15

than the former. The two maxima 2.86 and 3.01  $\mu$  of the band for water at 0° C. are much more pronounced in its intensity curve which is represented by the discontinuous line. The positions of the maxima are considerably shifted with respect to those for ice. The band for water at 4° C., represented by alternate dots and dashes is similar to that at 0° C., with the maximum shifted back to the long wave-length side to 2.93 and 3.05  $\mu$ . The right component

Table V.—Intensity Distribution of the H<sub>2</sub>O Band of Water at 38° C.

$\lambda$ in A.U.	$\nu$ in cm. <sup>-1</sup> .	$\delta\nu$ in cm. <sup>-1</sup> .	$\lambda$ I.R. in $\mu$ .	Intensity.
4790	20871	3834	2.61	6
4770	20959	3746	2.67	11
4748	21056	3649	2.74	24
4713	21212	3493	2.86	48
4687	21330	3375	2.96	31.5
4673	21394	3311	3.02	30
4658	21462	3243	3.08	24
4637	21560	3145	3.18	16.5
4617	21653	3052	3.28	2.2

Table VI.—Intensity Distribution of the H<sub>2</sub>O Band of Water at 98° C.

$\lambda$ in A.U.	$\nu$ in cm. <sup>-1</sup> .	$\delta\nu$ in cm. <sup>-1</sup> .	$\lambda$ I.R. in $\mu$ .	Intensity.
4813	20771	3934	2.54	1.5
4792	20862	3843	2.60	3.4
4760	21003	3702	2.70	11.6
4735	21113	3592	2.78	33.6
4707	21239	3466	2.89	48
4688	21325	3380	2.96	31.5
4669	21412	3293	3.04	23
4648	21509	3196	3.13	12.8
4625	21616	3089	3.24	2.6

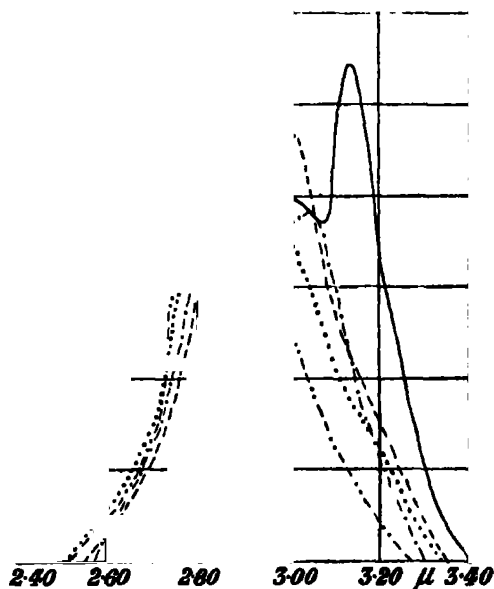


FIG. 3.—Intensity curves for Raman bands of ice and water at different temperatures. Ice —; water at °C., 0° — — —; 4°, — · — · —; 38°, . . .; 98°, — · — · —

is, however, much smaller in intensity than the left. At  $38^{\circ}\text{C}$ ., the water band given by the dotted line is not resolved at all, showing only a single maximum at  $2.89\ \mu$ . The asymmetry of the curve, however, indicates the presence of a second component. The curve for water at  $98^{\circ}\text{C}$ . is represented by the line drawn with two dots followed by two dashes. It reveals only one maximum at  $2.89\ \mu$  and the asymmetry is less marked than for the  $38^{\circ}\text{C}$ . band.

#### 4. *Explanation of the above Results.*

To explain the above changes in frequency and intensity of the water bands in different states and at different temperatures it is assumed that there are three types of molecules arising out of the association of the single molecules, viz.,  $(\text{H}_2\text{O})$ ,  $(\text{H}_2\text{O})_2$ , and  $(\text{H}_2\text{O})_3$ . These three types of molecules are supposed to give rise to the three components of the Raman band of water with frequencies about 3582, 3435, and  $3216\text{ cm.}^{-1}$  respectively, and the changes in their relative proportions with change of state and temperature are supposed to bring about the prominent intensity changes of the band.

In a recent communication, Sutherland\* suggested that it seems more satisfactory to assume that only two kinds of molecules are present, viz.,  $(\text{H}_2\text{O})$  and  $(\text{H}_2\text{O})_2$ . He attributed the line near  $3200\text{ cm.}^{-1}$  to the  $(\text{H}_2\text{O})_2$  molecules and explained the other two frequencies as arising out of a doubling of the  $\text{H}_2\text{O}$  frequency due to a condition of resonance degeneracy as has been observed in  $\text{CO}_2$ . In the latter case, the analysis of the bands is well known to permit of the application of the theory of perturbations. But for water the band analysis is very incomplete, and hence it is very difficult to say how far the conditions necessary for degeneracy are fulfilled in this substance. Dennison† has found that the perturbations which appear in  $\text{CO}_2$  do not occur in  $\text{CS}_2$  which is a molecule similar to it. As such the explanation that the two components in the water band arise out of a doubling due to degeneracy seems to be far fetched, when there is a lot of evidence to support the view that these components owe their origin to three different types of molecules.

The other argument put forward by Sutherland against the author's theory of the existence of three types of associated molecules in water is the change in frequency and structure of the Raman bands for ice from  $0^{\circ}\text{C}$ . to liquid air temperature. The bands, according to him, are much sharper at the lower temperature and have smaller Raman frequencies. Sutherland says that the

\* 'Proc. Roy. Soc.,' A, vol. 141, p. 542 (1933).

† 'Phys. Rev.,' vol. 41, p. 310 (1932).

frequency change is not to be expected according to the author's theory, as there cannot be any change in the structure of ice from  $0^{\circ}$  to liquid air temperature. He explains this change as being due to diminution of inter-molecular fields of force which arise owing to the highly polar nature of the  $\text{H}_2\text{O}$  molecules. The sharpening is attributed to the diminishing variations in these forces on account of the diminishing molecular rotations which bring about an increase in the field of force thus resulting in a diminution of frequency also.

Sutherland quotes Dennison's\* result that ice consists only of  $(\text{H}_2\text{O})_2$  molecules as evidence for his contention. But there does not seem to be any unanimity in the interpretation, by different authors, of the X-ray data of ice from which Dennison arrived at this result. While Dennison on the one hand, and Kinsey and Sponsler† on the other, infer that dihydrol is the only polymer occurring in ice, Bernal and Fowler‡ consider that there are three chief forms of arrangement of molecules of  $\text{H}_2\text{O}$  in water. They do not, however, admit of any structural analogy between these types of arrangement and the three associated molecules, their only objection to such an analogy being that the latter forms are conceived too much in the manner of molecular chemistry. It is sufficient therefore if it is mentioned that in view of the different interpretations of the X-ray results, it is not possible to arrive at any definite conclusions as regards the structure of ice and water.

If, according to Dennison, it is assumed that there are only the  $(\text{H}_2\text{O})_2$  molecules in ice, the presence of the second component at  $3135\text{ cm.}^{-1}$  in the Raman spectrum for ice obtained by Sutherland himself remains unexplained, except on the hypothesis that two types of molecules, viz.,  $(\text{H}_2\text{O})_2$  and  $(\text{H}_2\text{O})_3$ , exist even in this state.

One can very well understand the explanation given by Sutherland of the sharpening of the ice bands with diminishing temperature. But his interpretation of the frequency change as arising merely out of changes in inter-molecular fields of force does not seem to be tenable. The change with temperature in these fields in the same state cannot be of the same order of magnitude as that arising out of change of state. Hence the frequency changes in the two cases cannot be of the same order of magnitude as are actually observed.

A more probable explanation of the frequency shift which is consistent with the author's theory appears to be the following. For ice at  $0^{\circ}\text{ C.}$  the two

\* 'Phys. Rev.,' vol. 17, p. 20 (1921).

† 'Proc. Phys. Soc.,' vol. 45, p. 768 (1933).

‡ 'J. Chem. Phys.,' vol. 1, p. 521 (1933).

components are diffuse and unresolved from one another. On account of the overlapping of these, the resultant intensity distribution is considerably changed, so that the maxima of intensity are not in the same position as that corresponding to the sharp lines at liquid air temperature, but shifted with respect to the latter.

Sutherland cites the case of other polar molecules where there is a good deal of shift in Raman frequency with change of state, as supporting the view that the changes in the frequencies are due to changes in inter-molecular forces. But from chemical evidence, these substances are also known to be associated, and hence the explanation given by the author that the changes in frequency are due to association, applies equally well to these.

The following facts appear to favour the view that it is necessary to assume the presence of three types of molecules in water, viz.,  $(\text{H}_2\text{O})$ ,  $(\text{H}_2\text{O})_2$ , and  $(\text{H}_2\text{O})_3$  :—

- (1) In the liquid state there are three components of the band with Raman frequencies equal to 3216, 3435, and 3582  $\text{cm}^{-1}$ , the latter two being unresolved from one another.
- (2) Water-vapour gives rise to a single sharp Raman line with frequency 3650  $\text{cm}^{-1}$  which is nearly in the same position as the 3582 component in water, except for a slight shift of the latter due to the superposition of the next component over it.
- (3) Ice reveals only two components with Raman frequencies equal to 3196 and 3321  $\text{cm}^{-1}$  which are sharper than for water, but yet not completely resolved. These apparently correspond to the 3216 and 3435 frequencies of water.
- (4) There are systematic variations in the intensity of these bands with change of state and temperature which are not explicable except on the hypothesis of the existence of three types of molecules whose relative proportions vary with the above changes.
- (5) A confirmatory test of the hypothesis that the 3216 and 3435 components correspond to the triple and double molecules respectively is obtained in the work with water of crystallization in different substances, for which Nisi\* got the following Raman frequencies.

From Table VII, it is found that all substances which give rise to the band 3227  $\text{cm}^{-1}$ , which is attributed by the author to triple molecules, contain more than three  $\text{H}_2\text{O}$  molecules, indicating thereby the possibility of the forma-

\* 'Jap. J. Phys.,' vol. 7, p. 30 (1931).

tion of the triple molecules, and none of the substances containing less than this number give rise to this component.

The same argument applies to the allocation of the 3438 component to the double molecules, as all substances containing more than two molecules of water reveal this, the only exception to this rule being  $\text{Li}_2\text{SO}_4 \cdot \text{H}_2\text{O}$ , in which case it is possible that the  $\text{H}_2\text{O}$  molecule instead of being associated with another  $\text{H}_2\text{O}$  molecule is in combination as water of hydration either with the  $\text{Li}^+$  or  $\text{SO}_4^{--}$  ions which brings about the required change in frequency.

In spite of the fact that the nature and magnitude of the inter-molecular fields of force in crystals containing water of crystallization are different from those in ice, the frequencies of the  $\text{H}_2\text{O}$  bands in both are very nearly the same. This is an indication that they are not very much influenced by inter-molecular fields of force. Hence the only explanation of the frequency changes with change of state and of temperature of water appears to be that the nature of

Table VII.—Raman Frequencies of Water of Crystallization.

Substance.	Raman frequency in $\text{cm}^{-1}$ .		
$\text{Li}_2\text{SO}_4 \cdot \text{H}_2\text{O}$	—	3438	—
$\text{Na}_2\text{SO}_4 \cdot 10\text{H}_2\text{O}$	—	3438	—
$\text{AlK}(\text{SO}_4)_3 \cdot 12\text{H}_2\text{O}$	—	3384	—
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	3227	3445	—
$\text{CdSO}_4 \cdot \frac{1}{2}\text{H}_2\text{O}$	—	3423	—
$\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$	—	3404	3497
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	3211	3377	3404
$\text{Na}_2\text{S}_2\text{O}_3 \cdot 5\text{H}_2\text{O}$	3337	3414	—
$\text{C}_4\text{H}_8\text{O}_6\text{KNa} \cdot 4\text{H}_2\text{O}$	3303	3405	—

the water molecule itself changes, there being association of the single molecules into double and triple groupings.

According to the suggestion of Sutherland the two components 3435 and 3582 arise out of a doubling of the fundamental frequency of water due to resonance degeneracy. But in many substances containing water of crystallization only one component 3435 appears indicating the improbability of any such doubling.

#### *5. Relative Intensities of the Three Components of the Water Band as Determined from an Analysis of the Intensity Curves.*

Starting therefore with the hypothesis that the three components of the water band owe their origin to three different types of molecules viz.,  $\text{H}_2\text{O}$ ,

( $\text{H}_2\text{O}$ )<sub>2</sub>, and ( $\text{H}_2\text{O}$ )<sub>3</sub>, the intensity curves are analysed for a determination of the relative intensities of these components, from which it is proposed to evaluate the proportions in which these types exist in the different states of water.

The analysis is begun with the curve for ice. Three assumptions are made, without which it would not have been possible to proceed with the analysis. The first is that ice does not contain any single molecules and hence that its band consists of only two components corresponding to the double ( $\text{H}_2\text{O}$ )<sub>2</sub> and triple ( $\text{H}_2\text{O}$ )<sub>3</sub> groups.

The second is that, since in the resultant intensity curve the maximum on the longer wave-length side is much greater than the other, the wave-length of the maximum of the component on this side coincides with that of the above. The third assumption is that the component intensity curves are symmetrical about their central maximum.

The justification for the first assumption is this. Water-vapour is found to give a sharp line of Raman frequency  $\delta\nu = 3655 \text{ cm.}^{-1}$  corresponding to infra-red wave-length  $2.74 \mu$ . The presence of only a single line indicates that in this state there is only one type of molecules and that must be single  $\text{H}_2\text{O}$ . Since in the curve for ice there is very little intensity at this wave-length, it is natural to suppose that the component corresponding to single molecules is not present to contribute to the resultant intensity. The second assumption is just the converse of the fact that, if of two intensity curves superposed over one another giving rise to a common curve, one is much higher than the other, the larger maximum in the resultant curve very nearly coincides in position with that of the maximum of the larger component. The third assumption does not require any justification.

Water-vapour gives rise to a Raman line at  $2.74 \mu$ . In the liquid state the position of the maximum is at about  $2.9 \mu$ , and the band for ice consists of a larger maximum at  $3.13 \mu$ . The analysis reveals in general three components at  $2.73$ ,  $2.94$ , and  $3.13 \mu$  respectively. Since the first component corresponds to the Raman line for the vapour state, it will be called the vapour component and the single molecules which give rise to it will be called the vapour molecules. The second component should correspond to the higher polymer, namely, the double molecules, and since it is conspicuous in the liquid state, the molecules which give rise to it will be called the water molecules. The third component most conspicuous in ice is supposed to arise out of the triple molecules which on account of their preponderance in this state, will be called the ice molecules.



On the basis of the above assumptions, the intensity curve for ice is analysed in the manner represented in fig. 4. On the assumption that the ice component will have its maximum in the same position as that of the larger maximum in the resultant curve, a line is drawn through the latter parallel to the ordinate. As it is assumed that the component intensity curves will be symmetrical about their maximum, this line is taken as the line of symmetry for the ice component. Since the contribution to the resultant intensity curve to the extreme right will be due only to this component to some extent, the extreme portion of the curve is also taken as a portion of the curve due to the

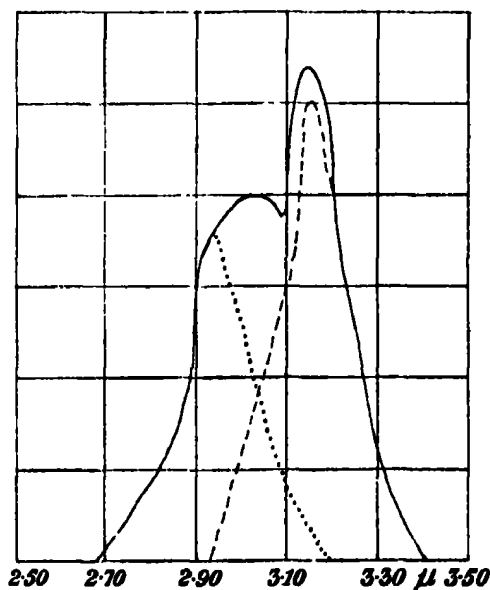


FIG. 4.—Analysis of the ice curve. Resultant curve —; ice component — — —; water component . . . . .

ice component. With this as the starting-point, a small curve is drawn on the left of the line of symmetry which appears as the inversion of the actual extreme portion of the resultant curve about this line. The curve so drawn is smoothly extended to meet the line of symmetry at some point. With this extended curve as the left half of the actual ice component, the right portion is completed. This complete curve which is drawn as discontinuous in the figure is taken as representing the curve for the ice molecules.

The rest of the analysis is simple. The ordinate differences between this curve and the resultant are taken corresponding to different abscissæ and plotted as a separate curve with the same abscissæ and the differences as

ordinates. The result is represented by the dotted curve, fig. 4, which is the second component corresponding to the  $(\text{H}_2\text{O})_2$  or water molecules.

The two components thus obtained are taken as those due to the ice and water molecules respectively and with the assumption that their shape, position, and extent should be the same for all states; the resultant curves for water at different temperatures are also analysed. For water at  $0^\circ \text{C}$ . of which the analysis is represented in fig. 5, a line is drawn parallel to the ordinate

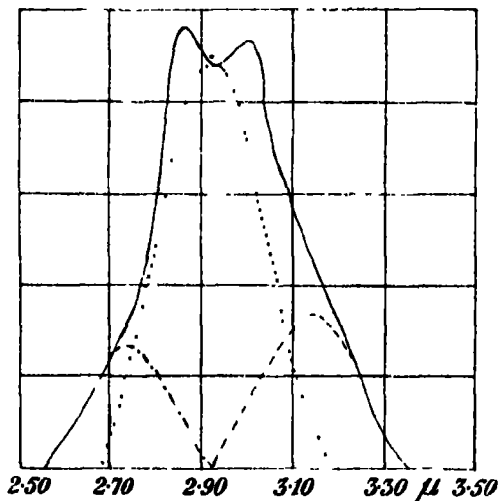


FIG. 5.—Analysis of the curve for water at  $0^\circ \text{C}$ . Resultant curve —; ice component — — —; water component . . . .; vapour component — . — . —.

with the abscissa equal to  $3.13 \mu$ , which corresponds to the maximum of the ice component. Another line parallel to the above is drawn through  $3.20 \mu$ , the extreme end of the water component to meet the resultant curve at some point. The portion of the resultant intensity curve lying between this point and its extreme right end is entirely due to the ice component. With the first line as the line of symmetry, a curve is drawn to its left, which is just the inversion of the above portion of the resultant curve about this line. It is smoothly extended to meet the line of symmetry and the complete curve is now drawn.

To get the water component, a third line is drawn also parallel to the ordinate with its abscissa equal to  $2.94 \mu$  which corresponds to the maximum of the water component. The ordinate differences between the resultant curve and the ice component, completed above, are plotted against their corresponding abscissæ and a portion of the water component thus obtained. The inversion of this curve about the line of symmetry is also drawn and produced to meet

this line at some point. With this point as the maximum of the water component in the curve for water at 0°, the whole component is completed. On subtracting the intensities of the above two components from the resultant, the ordinates for the third component are obtained and the curve corresponding to it is easily drawn.

From this analysis, the position, shape, and extent of all the three components are obtained and hence the analysis of the curves for water at the other temperatures is easily done.

Figs. 6, 7, and 8 respectively contain the results of the analysis of the intensity curves for water at 4°, 38°, and 98° C. From these, the positions of the maxima corresponding to the three types of molecules are easily deduced.

Another method of analysis is to start with water at 98° C. on the assumption that the position of the maximum for the vapour molecules is 2.74  $\mu$ , as in the Raman spectrum for water-vapour, and that the extreme portion of the curve to the left is entirely due to the vapour component. But since the latter is not at all resolved from the rest of the band, as contrasted to the ice component in the curve for ice, this procedure is not considered to be as reliable as the former.

Table VIII contains the wave-lengths and intensities of the maxima of the three components of the band due to the single, double, and triple molecules respectively.

Table VIII.—Positions of the Three Components of the H<sub>2</sub>O Band in Different States.

State.	Vapour molecules.		Water molecules.		Ice molecules.	
	$\lambda$ , $\mu$ .	Intensity.	$\lambda$ , $\mu$ .	Intensity.	$\lambda$ , $\mu$ .	Intensity.
Ice . . . . .	$\mu$	—	$\mu$		$\mu$	
Water 0° C.	2.73	14	2.93	35	3.13	50
Water 4° C.	2.78	16	2.93	43.5	3.13	17.5
Water 38° C.	2.79	21.5	2.93	48	3.13	20
Water 98° C.	2.80	26	2.92	36.5	3.13	15
Vapour 120° C.	2.74	—	2.93	36.5	3.13	9
Mean $\lambda$	2.77	—	2.93	—	3.13	—

The three components have therefore mean infra-red wave-length at 2.77, 2.93, and 3.13  $\mu$  respectively. The variations in the relative intensities of the components from one state to another are very conspicuous. The vapour component which is absent in the ice band systematically increases in intensity

with increasing temperature. The water component, however, has a maximum at  $4^{\circ}\text{C}$ . The ice component gradually diminishes in intensity with increasing temperature, with a sudden drop from the solid to the liquid state.

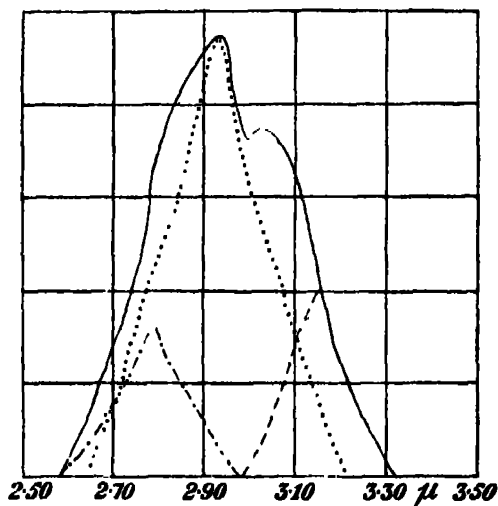


FIG. 6.—Analysis of the curve for water at  $4^{\circ}\text{C}$ .

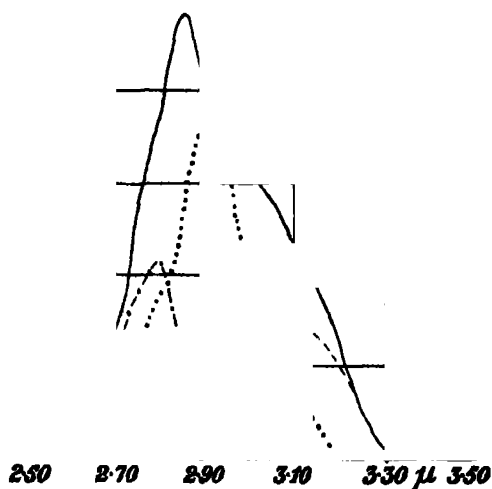


FIG. 7.—Analysis of the curve for water at  $38^{\circ}\text{C}$ .

#### 6. *Relative Proportions of the Three Types of Molecules.*

From the relative intensities of the three components, it is now possible to calculate the proportions of the ice, water, and vapour molecules at the different

temperatures. In this calculation it is assumed that the intensities are proportional to the number of molecules or groups of molecules and that the contribution to the intensity by a single molecule of  $\text{H}_2\text{O}$  is the same as that of  $(\text{H}_2\text{O})_2$  or  $(\text{H}_2\text{O})_3$ . The argument for the latter assumption is that a quantum of light incident upon a group of molecules can excite either the internal oscillations of one of the constituent molecules or the external oscillations between one molecule and another of the same group, but cannot excite the internal oscillations of all the molecules lying in a single group. So far, only Seagre\* has reported having observed a Raman band with  $\delta\nu = 140 \text{ cm.}^{-1}$  by excitation with the 2537 line of the mercury arc. It belongs to the external

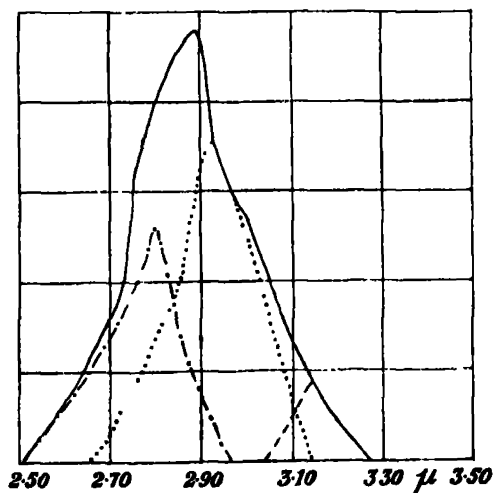


FIG. 8.—Analysis of the curve for water at 98° C.

FIGS. 6-8.—Resultant curve —; ice component — — — —; water component . . . .; vapour component — . . — .

oscillations of the polymerized molecules of water. It is exceedingly feeble compared to the band of water which has been analysed in this investigation, and can therefore be neglected. Hence, it can be assumed safely that the incident light energy is distributed among the three types of molecules according to their numbers. If it be assumed that the probability of excitation of all the three types of molecules is the same, the second assumption, viz., that the contribution to the intensity by a single molecule is the same as that of the double, which is also the same as that of the triple molecules, becomes valid.

\* 'Atti. Accad. Lincei,' vol. 13, p. 929 (1931).

On the basis of the above assumptions, the relative proportions of the three types of molecules, calculated from the intensities, are given in Table IX.

Table IX.—Relative Proportions of the Ice, Water, and Vapour Molecules in  $H_2O$  at different States.

Substance.	Intensity.			Percentage proportion of—		
	$H_2O$ .	$(H_2O)_2$ .	$(H_2O)_3$ .	$H_2O$ .	$(H_2O)_2$ .	$(H_2O)_3$ .
Ice	0	35	50	0 (0)	41 (32)	59 (68)
Water 0° C.	14	37·5	17·5	19 (9)	58 (57)	23 (34)
Water 4° C.	16	48	18	19·5 (10)	58·5 (58)	22 (32)
Water 38° C.	21·5	36·5	15	29 (16)	50 (52)	21 (32)
Water 98° C.	26	36·5	9	36 (21)	51 (58)	13 (21)

In the first column is given the state of water. The second, third, and fourth columns contain the intensities of the single, double, and triple molecules respectively. In the last three columns are given the percentage relative proportions of the three types of molecules. The top numbers give their proportion in numbers and the lower ones within brackets the proportion by weight.

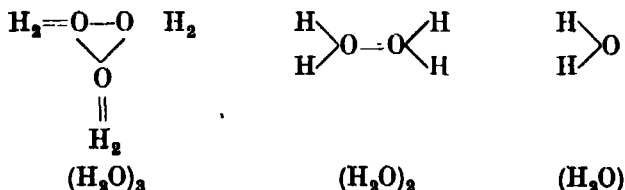
The general results of the above table are :—

- (1) The number of single or vapour molecules progressively increases in intensity with increasing temperature.
- (2) The number of double or water molecules remains practically constant except in the transition from ice to water at 0° C. where there is a sudden increase. They appear to be maximum at 4° C.
- (3) The triple or ice molecules suddenly diminish in number from ice to water at 0° C. and then gradually diminish with increasing temperature.

### 7. Significance of the Results.

The sudden diminution in the number of triple molecules from ice to water seems to be connected with the sudden change in specific volume or density

which is so well known. This is probably due to the larger volume occupied by the triple molecules than the double. If we assume the graphic formula for these molecules, suggested by Sutherland,\* the above result can be easily explained. The ice, water, and vapour molecules, which he calls trihydrol, dihydrol, and hydrol respectively, have the following formulæ:—



The trihydrol obviously occupies more than one and a half times the volume of the dihydrol. Hence in ice, which consists mostly of ice molecules, trihydrol, the packing is less compact than in water, which consists mostly of water molecules, dihydrol. The diminution in specific volume from ice to water is therefore a natural consequence of the change of ice molecules to water molecules. The maximum in the number of water molecules at 4° C. explains the minimum specific volume or maximum density of water at this temperature. This cannot be taken as a reliable result, because the intensity values have an error of about 10% and the variation in the number of double molecules from water at 0° C. to water at 4° C. is only about 2%.

An alternative and a more probable explanation of the maximum density of water at 4° C. seems to be as follows. There are two causes for the change in density. The first is due to thermal expansion of water with temperature which involves a diminution in density. The second is due to the change in the relative proportions of the three types of molecules. Since the double molecules are really constant in number, the change can be due to the dissociation of triple to single molecules which brings about a closer packing of the molecules and hence a higher value of the density. These two processes take place in opposite directions, and 4° C. is perhaps the optimum temperature at which these two effects are such that the packing is closest.

## 8. Discussion.

Many people have put forward the hypothesis of the polymerization of water into double and triple molecules. But the only attempt so far made to determine quantitatively the proportions of the different types in the

\* 'Phil. Mag.,' vol. 50, p. 460 (1900).

various states is that of Sutherland as early as 1900. The method adopted by him was from the density determination. The curve for the variation of the specific volume of water with temperature, which attains a minimum value at  $4^{\circ}\text{C}$ ., was made continuous by him and was extrapolated to the zero temperature, from which the density of dihydrol came out to be 1.089. He also got the value 0.88 for trihydrol, and from the density of water at different temperatures he calculated the relative proportions of these two types of molecules at the various temperatures. For water at  $0^{\circ}\text{C}$ . he obtained 37.5% of trihydrol and 62.5% of dihydrol. Compared to this, the values obtained by the author are 34 and 57 respectively together with 9% single molecules.

Sutherland assumed that there are no single molecules at all in water, which does not seem to be valid. But since water consists mostly of double molecules his values do not err greatly.

There is one advantage in the method adopted by the author. There is direct evidence of the presence of three types of molecules, single, double, and triple, or even still higher polymers. Since the number of molecules or aggregates of molecules is directly proportional to the intensity of the Raman band, we have in this a more direct method of estimating their percentages. The order of accuracy is the same as is available in the determination of the intensities of spectral lines.

### *Summary.*

The Raman bands for ice and water at  $0^{\circ}$ ,  $4^{\circ}$ ,  $38^{\circ}$ , and  $98^{\circ}\text{C}$ . are taken and their intensity distribution with respect to wave-length determined. The general appearance of the intensity curves is found to indicate that each of them is due to superposition of three components, which are attributed to the single, vapour molecules, double, water molecules, and triple, ice molecules, respectively. Each of the intensity curves are analysed into the three components and from the intensities of the components the relative proportions of the ice, water, and vapour molecules are calculated. The changes in  $\text{H}_2\text{O}$  from one state to another are explained from the above results.

---



## *The Specific Heat of Nickel and of some Nickel-Copper Alloys.*

By K. E. GREW, Ph.D.

(Communicated by R. Whiddington, F.R.S.—Received November 22, 1933.)

### *Introduction.*

The study of the alloys of the ferromagnetic with other elements promises to elucidate some of the problems of ferromagnetism. Sadron,\* for instance, has determined the "ferromagnetic moment" of certain foreign elements in nickel. He finds that elements in the same column of the periodic table have nearly the same moment, and that in passing across the table the value of this moment first increases and then decreases. Stoner† has recently suggested an interesting interpretation of these results in considering the moments of ferromagnetics in general.

The variation of the specific heat with temperature in such binary alloys apparently has not yet been investigated, although the specific heat of the ferromagnetics in the pure state has been the subject of several investigations.

As is well known, the general characteristics of the temperature dependence of the specific heat of ferromagnetics—an excess over the "normal" value for non-ferromagnetics increasing to a maximum as the temperature rises to the Curie point, followed by a more or less sudden decrease—are satisfactorily accounted for by the Weiss molecular field hypothesis which requires the introduction of a term  $S_M$  due to the negative magnetic energy :

$$S_M = \frac{1}{2 \cdot J} N \rho \frac{d\sigma^2}{dT},$$

where

$S_M$  = "magnetic" specific heat in cal./gm. deg.

$J$  = mechanical equivalent of heat

$N$  = the Weiss molecular field constant referred to unit volume

$\rho$  = density

$\sigma$  = spontaneous magnetization per unit mass.

Mme. Lapp‡ has confirmed the earlier observations of Weiss and his collaborators that in nickel the observed change at the Curie point is satisfactorily

\* 'Ann. physique,' vol. 17, p. 371 (1932).

† 'Phil. Mag.,' vol. 15, p. 1018 (1933).

‡ 'Ann. physique,' vol. 12, p. 442 (1929).

accounted for by the disappearance of the term  $S_M$ . She has shown, however, that after correcting the observed values for quantum variation, dilatation, and for the magnetic term  $S_M$ , the specific heat is still in excess of the "normal" value for a substance identical in every respect except that it has no intrinsic magnetization. The origin of this excess is unknown.

The purpose of the present work was to investigate the temperature dependence of the specific heat of some binary alloys of nickel, for which no data appears hitherto to have been available. The cupro-nickels were chosen as their magnetic properties have already been investigated by Alder.\* Measurements were also made on pure nickel, both as a check on the method and because of the disagreement near the Curie point among recent values.

### *Preparation of the Alloys.*

The alloys were prepared from pure nickel and copper supplied by Messrs. Hilger. The impurity in both metals was given as 0.03%. For the preparation of the alloys I am indebted to the National Physical Laboratory. Three alloys, of composition approximately 94, 87, 79% nickel, were made, using an induction furnace and an atmosphere of hydrogen. The mechanical properties of these alloys are poor, and the reduction of the ingots to rod form presented considerable difficulties, but these were successfully overcome, and rods of 2 mm. diameter and a metre or so in length were made.

After preparation the alloys were analysed. The composition so determined was :—

Alloy No.	Cu %.	Ni % (by difference).
1	5.99	94.01
2	12.77	87.23
3	21.19	78.81

I am indebted to the Department of Metallurgy, Yale University, for this analysis.

### *Experimental Method.*

The most recent measurements of the specific heat of pure nickel are those of Lapp (*loc. cit.*). In her method the specimen in the form of a rod is brought to a steady temperature. A current is then passed through it and the rise in

\* "Thesis, Zurich" (1916).

temperature due to its own resistance is measured by a thermocouple inserted in the rod. If a current  $I$  produces a temperature change  $\Delta T$  in time  $\Delta t$ , and if  $m$  is the mass of the rod and  $r$  its resistance, the mean specific heat between  $T$  and  $T + \Delta T$  is

$$s = \frac{I^2 r}{m \cdot J \cdot \frac{\Delta T}{\Delta t}}. \quad (1)$$

The satisfactory results thus obtained for pure nickel suggested the adoption of this method for the copper-nickel alloys.

The experimental arrangement did not differ essentially from that used by Lapp, except that, whereas Lapp used a specially designed wattmeter for the measurement of the energy supplied, in the present work the current and resistance were measured separately. The specimen was a rod of 2 mm. diameter and length 70 cm. It was mounted as shown in fig. 1, in which  $AB$  is the thermocouple spot-welded to the rod;  $CD$  are leads of the same material as the rod, drawn tightly through two holes in the rod; these served to measure the resistance of the intervening part, which had a length of some 15 cm. and a mass  $m$  determined from measurements of the mass per unit length made on a second similar rod;  $EF$  are the leads for the heating current  $I$ . The specimen was enclosed in a pyrex tube which could be evacuated in order to reduce the heat transfer between the specimen and the tube walls. The tube was inserted in a furnace in which the temperature gradient had been reduced to less than  $0.5^\circ \text{C}$ . over a length of 10 cm. at  $350^\circ \text{C}$ ., and to smaller values at lower temperatures. For measurements at low temperatures the tube was immersed in a bath of liquid air, or carbon dioxide snow in alcohol.

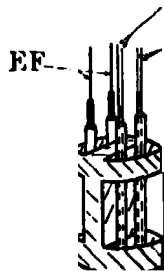


FIG. 1. — Specimen mounting and connections.

The heating current  $I$  varied from 3 to 5 amps. and was measured with a potentiometer and a standard resistance. The resistance  $r$  between the leads  $CD$  was determined from the potential difference due to the passage of a known current which was sufficiently small to give no appreciable heating effect. In calculating the energy, the resistance involved corresponds to the mean of the initial and final temperatures; this was taken from the resistance-temperature curve.

The temperature measurements were made with a platinum-platinum-rhodium couple for the high temperatures and a copper-constantan couple for the low temperatures. Both couples were calibrated by comparison with couples calibrated by the Bureau of Standards. The diameter of the wires was small (0.2 mm.) in order to minimize any disturbing effect at the point of attachment to the rod through thermal conduction or resistance change. In attaching the couple the two wires were bent at right angles to the rod, brought together from opposite sides, and then each welded to the rod. The e.m.f. was measured on the same potentiometer used for measuring the current. The tables given by Adams\* and by Southard and Andrews† were used in deducing the temperature.

In determining the rate of increase of temperature the procedure was as follows. When the rod was in temperature equilibrium with the furnace, the e.m.f. of the couple was balanced across the potentiometer. The heating current, which had been flowing through a subsidiary circuit for some time to attain a steady value, was then passed through the specimen, and the galvanometer used in conjunction with the potentiometer deflected as the temperature of the rod rose. The time of passage of the spot across given scale divisions was recorded on a chronometer. When the temperature had risen 2–4°, the heating current was broken and the recording continued for a time as the rod cooled. The current was then again passed through the rod and its value measured. From the chronometer record the heating curve and the initial part of the cooling curve were plotted, and a correction for radiation loss was deduced in the usual manner. The corrected heating curve was always accurately linear, and from it the rate of deflection  $\Delta\theta/\Delta t$  was found. The rate of rise of temperature could then be calculated from this, the measured sensitivity of the galvanometer  $\Delta\theta/\Delta E$ , and the slope  $\Delta E/\Delta T$  of the calibration curve of the couple at the temperature  $T'$  which is the mean of the initial and final temperatures. All the quantities in the expression (1) for the mean specific heat over the range  $T, T + \Delta T$  were then known.

The small e.m.f. introduced into the thermocouple circuit by the passage of the heating current was compensated, as in Lapp's experiment, by the introduction of an opposing e.m.f. The sensitivity of the galvanometer was determined immediately before the passage of the heating current by setting up a known small e.m.f. in the thermocouple circuit.

\* 'J. Amer. Chem. Soc.,' vol. 36, p. 65 (1914).

† 'J. Franklin Inst.,' vol. 207, p. 323 (1929).

*The Precision of the Measurements.*—Lapp has given an exhaustive account of the errors in her method. The present method differs from hers only in detail, and the error in the final value of the specific heat at a given temperature remains approximately the same, *i.e.*, between 1.5 and 2%.

### Experimental Results.

The curves of the specific heat as a function of the temperature are shown in figs. 2 and 3, and the numerical data are given in Tables I and II. The

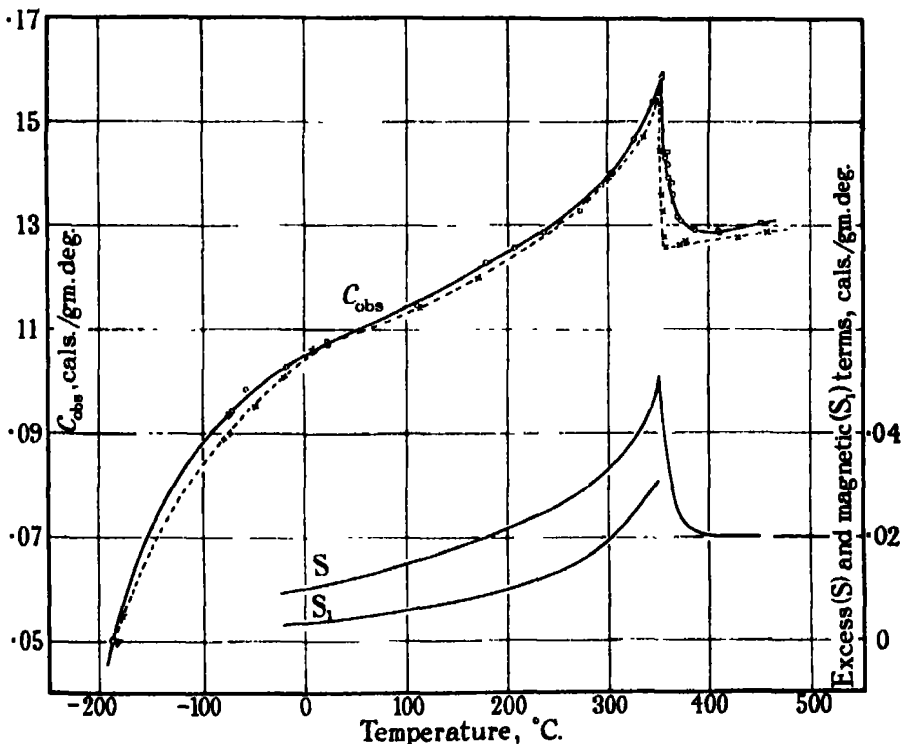


FIG. 2.—Pure nickel.  $C_{obs}$ . —○— observed values of specific heat; —x— Lapp's values; —v— Rodebush and Michaleck's values.  $S$ , excess specific heat;  $S_1$ , magnetic specific heat.

results obtained by Lapp\* for pure nickel are indicated by the broken line in fig. 2. It will be seen that the present results are generally higher than hers, but the difference is of the magnitude of the experimental error except in the low temperature readings, and here the present measurements agree well with those of Rodebush and Michaleck.†

\* 'Ann. Physique,' vol. 12, p. 442 (1929).

† 'J. Amer. Chem. Soc.,' vol. 47, p. 2117 (1925).

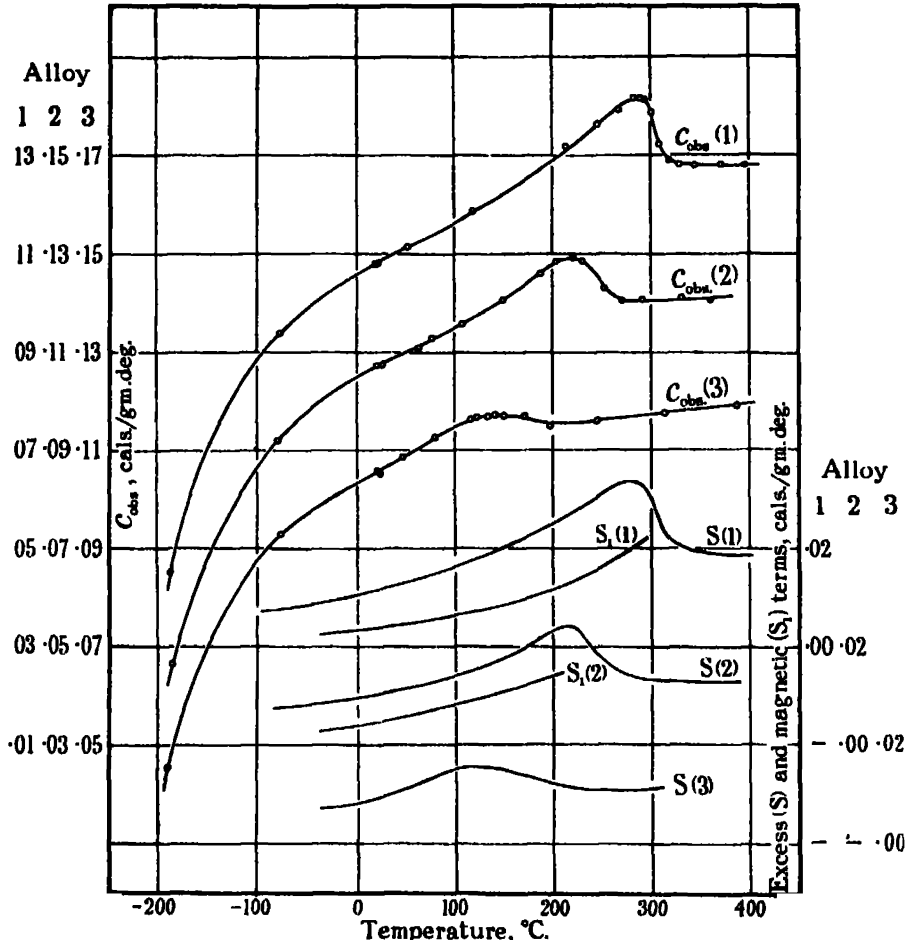


FIG. 3.—The alloys. (1) 94.0% ; (2) 87.2% ; (3) 78.8%.  $c_{obs}$ , observed values of specific heat ; S, excess specific heat ;  $S_1$ , magneto specific heat.

Table I.—The Specific Heat  $c_{obs}$  cal./gm. of Nickel.

T °C.	$c_{obs}$ .	T °C.	$c_{obs}$	T °C.	$c_{obs}$ .
22.2	0.1077	328.1	0.1465	377.7	0.1330
22.3	0.1078	345.5	0.1537	388.0	0.1292
22.9	0.1078	352.1	0.1557	410.7	0.1288
23.7	0.1070	354.2	0.1587	410.9	0.1286
111.8	0.1148	358.2	0.1432	453.0	0.1307
179.0	0.1230	360.8	0.1417	Low temperature measurements.	
206.7	0.1258	361.5	0.1442		
237.0	0.1290	361.8	0.1392		
273.1	0.1332	364.9	0.1384		
295.4	0.1380	367.4	0.1360		
305.4	0.1402	370.0	0.1316	20.2	0.1077
325.8	0.1465	373.0	0.1311	17.4	0.103
				56.9	0.099
				71.0	0.094
				187.0	0.050

Table II.—The Specific Heat  $c_{obs.}$  cal./gm. of Nickel-copper Alloys. Alloy 1, 94.0% Ni; alloy 2, 87.2% Ni; alloy 3, 78.8% Ni.

Alloy 1.		Alloy 2.		Alloy 3.	
T °C.	$c_{obs.}$	T °C.	$c_{obs.}$	T °C.	$c_{obs.}$
23.4	0.1081	27.3	0.1075	24.3	0.1056
52.4	0.1115	83.8	0.1107	47.9	0.1088
118.3	0.1186	76.8	0.1128	81.4	0.1130
213.4	0.1318	103.0	0.1148	117.5	0.1167
246.4	0.1364	148.1	0.1202	123.7	0.1172
268.4	0.1391	148.9	0.1205	134.5	0.1171
283.3	0.1416	187.8	0.1256	141.6	0.1178
288.6	0.1416	205.8	0.1286	149.5	0.1174
293.9	0.1413	220.5	0.1289	171.5	0.1174
300.2	0.1388	232.2	0.1284	197.1	0.1155
309.0	0.1322	253.1	0.1228	245.3	0.1166
319.6	0.1293	270.7	0.1205	313.9	0.1182
323.1	0.1282	291.9	0.1208	355.8	0.1202
345.3	0.1278	326.7	0.1212	386.8	0.1192
382.2	0.1279	360.7	0.1206		
396.9	0.1276				
20.4	0.1080	20.7	0.1075	21.4	0.1061
-76.6	0.094	-78.6	0.092	-76.0	0.094
-190.0	0.046	-188.3	0.047	-191.9	0.046

Besides the measurements of Lapp on pure nickel there are those of Potter and Sucksmith\* and Klinkhardt.† All observers agree in finding a remarkably abrupt change as the Curie point is approached from the low temperature side. Immediately above the Curie point there is disagreement, however. Lapp finds that the change extends over only 5–7° before the steadily increasing values are reached. Potter and Sucksmith find a change extending over 30–40°; moreover, their values above the Curie point are some 5% higher than those of Lapp. Klinkhardt's measurements show an extended change and values above the Curie point rather lower than Lapp's. The present measurements indicate an extended change over 40° and confirm the values of Lapp at the higher temperatures. This difference in behaviour is too large to be attributed to experimental error: the cause lies possibly in dissimilar treatment of the specimens.

In general, the addition of copper, besides lowering the Curie point, extends the change associated with the transition from the ferromagnetic to the paramagnetic state over a greater temperature range, and decreases its magnitude.

\* 'Proc. Roy. Soc.,' A, vol. 112, p. 157 (1928).

† 'Ann. Physik,' vol. 84, p. 167 (1927).

*Discussion of Results.*

*The "Excess" Specific Heat Term.*—The observed specific heat  $c_{\text{obs}}$  at constant pressure may be regarded as composed of the following terms :

$$c_{\text{obs.}} = c_v + (c_p - c_v) + c_i + S,$$

where  $c_v$  is the specific heat at constant volume according to the Debye theory ;  $(c_p - c_v)$  is the correction for dilatation ;  $c_i$  represents the excess over the Debye value common to most metals ; and the "excess term"  $S$  includes effects due to the magnetization and other causes.

(a) *The Quantum Variation.*—The quantity  $c_v$  can be determined from the Debye expression, which gives the specific heat as a function of  $T/\theta$  where  $\theta$  is the characteristic temperature of the substance. As Lapp has shown, the characteristic temperature is best determined from a measurement of the specific heat at a low temperature. Now the results for pure nickel and the alloys at the temperature of liquid air suggest that at this temperature the effect of the added copper is small. Hence, although in determining the characteristic temperature  $\theta$  the observed values of the specific heat have to be corrected for dilatation and the magnetic term, the corrections are small, and it may be assumed that the characteristic temperature, and thus the quantum variation, is the same for the alloys as for pure nickel. The present values of the specific heat of pure nickel at the temperature of liquid air are somewhat higher than those of Lapp and of Rodebush and Michaleck, which are here in good agreement. The discrepancy is to be attributed to the imperfect calibration of the couple used in the present work at this low temperature. Lapp's values for the specific heat at this temperature and for the resulting quantum variation have therefore been used. The error in the excess specific heat of the alloys arising from the assumption that the quantum variation is the same as for pure nickel is estimated at 1%.

(b) *The Dilatation Correction.*—The difference between the specific heats at constant pressure and at constant volume is given by

$$(c_p - c_v) = \frac{T}{J} \frac{\alpha^2}{\rho\gamma},$$

where

$T$  = absolute temperature,

$J$  = mechanical equivalent of heat,

$\rho$  = density,

$\alpha, \gamma$  = coefficients of cubical expansion and of compressibility respectively.



A knowledge of the temperature variation of the coefficients of expansion and of compressibility appropriate to each alloy would therefore be required for a precise evaluation of the correction for dilatation. However, using what values are available for these coefficients at various temperatures for pure nickel and copper, it appears that except in the neighbourhood of the Curie point (where the coefficient of expansion of nickel shows an abrupt change) the corrections for the alloys may be taken as the same as for pure nickel without introducing an error into the excess specific heat  $S$  of more than 1%. This assumes that the coefficients vary approximately linearly with composition: the results of Krupkowski\* on the coefficient of expansion and of Johansson† on the elastic modulus support this assumption. Lapp's values of  $(c_p - c_v)$  have therefore been applied to the alloys as to pure nickel, except in the neighbourhood of their Curie points. The abrupt change at the Curie point in the correction for pure nickel will presumably be less in the alloys. It has been calculated on the assumption that its magnitude varies with composition in the same way as the saturation moment, *i.e.*, that it varies linearly with the composition and becomes zero for an alloy of composition approximately 40% nickel. The values of  $(c_p - c_v)$  immediately below the Curie point can then be found by a somewhat arbitrary graphical interpolation, which, however, introduces little error.

(c) *The Small Excess over the Theoretical Value.*—In correcting for this, Lapp assumed that the correction was the same for nickel as for copper, and the assumption may be extended to the alloys.

Making these three corrections to the observed value of the specific heat gives the excess term  $S$ , of which the variation with temperature for the three alloys and for pure nickel is shown in figs. 2 and 3, and in Table III.

This excess term may be regarded as composed of two parts: the first part  $S_1$  is attributed to the change in intrinsic magnetization  $\sigma$ ; the other part  $S_2$  is simply that part of  $S$  not accounted for by  $S_1$  and must be regarded as of unknown origin.

*The Magnetization of the Alloys.*—On the Weiss hypothesis the specific heat of a ferromagnetic includes a term

$$S_1 = -\frac{1}{2} N \rho \frac{d\sigma^2}{dT}, \quad (2)$$

where  $\sigma$  is the specific magnetization and  $\rho$  the density.

\* 'Rev. Métall.', vol. 26, p. 131 (1929).

† 'Ann. Physik,' vol. 76, p. 445 (1925).

To calculate  $S_1$  a knowledge of the temperature variation of the magnetization is required. For this recourse was had to the work of Alder\* who gives values for the spontaneous magnetization as a function of the temperature for nickel-copper alloys of various compositions. (Some typical curves are given in the 'International Critical Tables,' vol. 6, p. 405.) By interpolation from his results magnetization-temperature curves applicable to alloys of composition

Table III.—The Excess Specific Heat Term  $S$  cal./gm. deg. Alloy 1, 94.0% Ni; alloy 2, 87.2% Ni; alloy 3, 78.8% Ni.

T ° C.	Excess specific heat, S cal./gm.		T ° C.	Excess specific heat, S cal./gm.	
	Nickel.	Alloy 1.		Alloy 2.	Alloy 3.
-100	0.0074	0.0084	-50	0.0082	0.0078
0	0.0107	0.0112	0	0.0103	0.0090
100	0.0150	0.0170	25	0.0115	0.0101
200	0.0216	0.0257	50	0.0125	0.0116
225	0.0236	0.0295	75	0.0139	0.0137
250	0.0258	0.0313	100	0.0150	0.0156
275	0.0284	0.0343	125	0.0166	0.0163
300	0.0323	0.0316	150	0.0185	0.0153
325	0.0379	0.0212	175	0.0210	0.0144
350	0.0475	0.0203	200	0.0238	0.0127
375	0.0225	0.0195	225	0.0242	0.0120
400	0.0200	0.0180	250	0.0186	0.0118
425	0.0200		275	0.0147	0.0115
			300	0.0144	0.0117

the same as those used in the specific heat measurements were drawn. From these curves the value of  $d\sigma^2/dT$  was found, though not with much accuracy; it would have been preferable to follow Lapp in deducing it from magneto-caloric data had such data been available.

Alder investigated also the susceptibility of the alloys in the paramagnetic state. His results showed that the molecular field constant was approximately the same for all alloys containing more than 50% nickel, and had a value  $N\rho \sim 117000$ .

It is interesting to note that the value of the Curie point deduced by interpolation from Alder's results agrees well, as Table IV shows, with the value found from the resistance-temperature curves obtained as part of the specific heat measurements. These curves show, at the Curie point, a change of slope which, unlike the specific heat change, does not extend over an increasing temperature range as the nickel content diminishes.

\* "Thesis, Zurich" (1916).

The agreement between the values deduced from Alder's measurements and those found in the present work suggests that Alder's results may be applied with some confidence.

*The Molecular Field Constant and the Evaluation of the Magnetic Term.*—The calculation of the magnetic term in the specific heat from the expression (2) involves a knowledge of the molecular field constant  $N$  appropriate to the ferromagnetic state. There is unfortunately no means of determining this constant directly.

Table IV.

Alloy % Ni.	Saturation moment.	Curie point ° C. (magnetization).	Curie point ° C. (resistance).
100.0	57.2	358	359
94.0	51.6	301	302
87.2	45.3	240	238
78.8	37.3	163	159

At first sight a possible procedure would appear to be as follows. Essentially the value of  $N$  is derived from the relation

$$\theta = \frac{N\rho\sigma_M^2}{3MR} \cdot \frac{j+1}{j},$$

where  $\theta$  is the Curie point,  $M$  is the molecular weight,  $\rho$  is the density,  $\sigma_M$  is the moment per mole, and  $j$  is the mechanical moment of the carriers. But

$$\sigma_M^2 \cdot \frac{j+1}{j} = p^2 M_w^2,*$$

where  $p$  is the magneton number defined by the relation

$$p = \frac{1}{M_w} \sqrt{3R\chi_M (T - \theta)},$$

in which  $\chi_M$  is the molar susceptibility and  $M_w$  is the Weiss magneton. Thus

$$\theta = \frac{N\rho p^2}{3MR} M_w^2.$$

In the paramagnetic state the value of  $p$  can of course be directly determined. In the ferromagnetic state  $p$  can be found only indirectly from the atomic moment and the  $j$  value. The magnetization-temperature curves for nickel

\* cf. Stoner, 'Phil. Mag.,' vol. 10, p. 34 (1930).

are consistent with a value of  $j = 1/2$ , and the low temperature saturation moment is 3 magnetons. The magneton number appropriate to the ferromagnetic state is therefore

$$p = 3\sqrt{3} = 5.2,$$

as compared with  $p = 8.0$  in the paramagnetic state. This gives  $Np > 272000$  in the ferromagnetic state (the inequality sign allowing for possible  $j$  values other than  $1/2$ ) and  $Np = 117000$  in the paramagnetic state.

But if the value of the magnetic term  $S_1$  be calculated, using the value  $Np = 272000$ , and values of  $d\sigma^2/dT$  determined experimentally, then at temperatures above  $150^\circ \text{C}$ .  $S_1$  is greater than the observed excess specific heat  $S$ . This method of calculating the molecular field constant must therefore be inapplicable; some factor not considered in the quantized Weiss theory appears to be involved.

The value of  $N$  in the ferromagnetic state therefore cannot be directly ascertained. It is known, however, that the observed change in the specific heat of nickel at the Curie point is accounted for quantitatively by the disappearance of the magnetic term  $S_1$ , when in the calculation of  $S_1$  the value of  $N$  appropriate to the paramagnetic state is used. In the absence of further knowledge it is not unreasonable to assume, following Weiss and Lapp, that this value of  $N$  is applicable not only at the Curie point but below this temperature. Only part of the excess specific heat  $S$  is thus accounted for by the magnetic term  $S_1$ . The difference between  $S$  and  $S_1$  is represented by another term  $S_2$ , the origin of which is unknown.

The curves of figs. 2 and 3 showing  $S_1$  as a function of temperature have been calculated on this basis.

*The Change in Specific Heat at the Curie Point.*—At the Curie point the specific heat  $S_1$  due to the intrinsic magnetization reaches a maximum and then disappears. Weiss and Beck,\* and later Lapp,† have shown for pure nickel that the observed change in specific heat agrees closely with the calculated maximum value of  $S_1$ . It is interesting to compare these quantities in the case of the alloys. In Table V the values of  $S_1$  at the Curie point and the change  $\Delta S$  in the excess specific heat (determined by extrapolation from above and below the Curie point) are shown.

Remembering that the determination of both  $S_1$  and  $\Delta S$  involves a very uncertain extrapolation to the Curie point (especially for the 78.8% alloy) and that the error in  $\Delta S$  alone is of the order of 10%, the agreement is as close

\* 'J. Physique,' vol. 7, p. 249 (1908).

† 'Ann. Physique,' vol. 12, p. 442 (1929).

as could be expected. Thus, in the alloys as in pure nickel, the change in specific heat at the Curie point can be explained quantitatively by the disappearance of the magnetic term (calculated on the assumption that the value of the molecular field constant deduced from the paramagnetic behaviour is applicable at this temperature).

*The Term  $S_2$  of Unknown Origin.*—Because of the uncertainty as to the appropriate value of the molecular field constant below the Curie point it is undesirable to attach too much importance to the division of the excess specific heat into two parts,  $S_1$  and  $S_2$ . But if the idea be retained for convenience,

Table V.

Alloy % Ni.	$S_1$ cals./gm. deg.	$\Delta S$ cals./gm. deg.
100.0	0.033	0.034
94.0	0.023	0.020
87.2	0.017	0.015
78.8	(0.009)	0.010

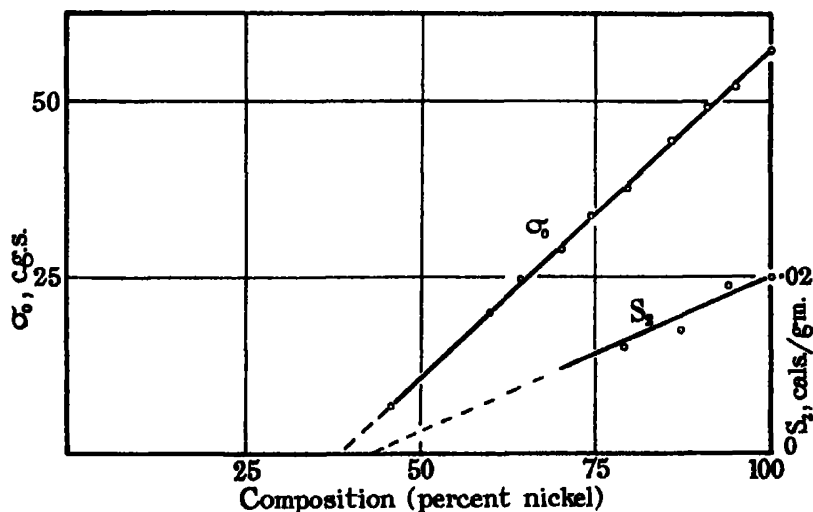


FIG. 4.—The variation with composition of the saturation moment  $\alpha_0$  and of the second, "unknown," term  $S_2$ .

then  $S_2$  must be regarded as increasing as the temperature approaches the Curie point, ultimately becoming equal to the excess specific heat above the Curie point since  $S_1$  is here absent.

In fig. 4 the value of  $S$  or  $S_2$  above the Curie point for the different alloys is shown as a function of the composition. The low temperature saturation

moment is also shown. Although the evidence is far from conclusive, it appears that  $S_2$  and the saturation moment may both become zero at the same value of the composition. This suggests that  $S_2$  and the magnetization may have a common origin.

This work was done at Yale University during the tenure of a Commonwealth Fund Fellowship. It is a pleasure to thank Professor McKeehan for his advice and for the provision of every facility, and Dr. Stoner for his kindly criticism.

### *Summary.*

The specific heat of pure nickel and of three alloys of copper with nickel (94.0, 87.2, 78.8% nickel respectively) has been measured over a temperature range from  $-180^\circ$  to  $450^\circ$  C. An electrical method was used in which the specimen was heated by the effect of its own resistance, the temperature change being measured by means of a thermocouple attached to the specimen.

The experimental results are given in Tables I and II and are shown graphically in figs. 2 and 3.

The observed values for the alloys were corrected (as in Lapp's work\* on pure nickel) for quantum variation, dilatation, etc., and the excess  $S$  over the "normal" value for a substance without intrinsic magnetization but otherwise similar was deduced. The specific heat  $S_1$  due to the intrinsic magnetization was calculated, using the magnetization measurements of Alder.† The value of the molecular field constant  $N$  which is involved in the calculation was taken to be the same as that determined from susceptibility measurements in the paramagnetic state. The existence of a second term  $S_2$  in the excess specific heat  $S$  must then be assumed, since  $S_1$  accounts for only a part of  $S$ .

Although the values of  $S$  and  $S_1$  are not known with much precision, it appears that the change in  $S$  at the Curie point in the alloys can be explained quantitatively by the disappearance of the term  $S_1$ , as for pure nickel and iron, Table V.

The value of  $S$  above the Curie point, regarded as a function of the composition of the alloys, probably becomes zero when the saturation moment becomes zero, fig. 4. This suggests that the second term  $S_2$  has a common origin with the ferromagnetism.

\* 'Ann. Physique,' vol. 12, p. 442 (1929).

† "Thesis, Zurich" (1916).

## *A Theory of the Electrical Breakdown of Solid Dielectrics.*

CLARENCE ZENER, H. H. Wills Physics Laboratory, Bristol.

(Communicated by R. H. Fowler, F.R.S.—Received December 27, 1933, Revised March 1, 1934.)

### § 1. *Introduction.*

In the modern theory of metallic conduction initiated by Bloch, conductors, semi-conductors, and non-conducting crystals may be represented by the same model. In this model each electron is supposed to move freely in the periodic field of the lattice. Owing to this field, not all electronic energy levels are allowed; the allowed levels are grouped into bands, separated by energy intervals which are disallowed. If all the energy levels of a given band are occupied by electrons, then, according to the theory, these electrons can make no contribution to an electric current in the crystal. If all the bands are full, the crystal must be an insulator. Thus in an insulator there exist a number of energy bands which are completely full, and a number of bands of higher energy which, for a perfect crystal at the absolute zero of temperature, are empty.

In a real non-conducting crystal, however, there will be a few electrons in the first unfilled band, owing to thermal excitation, impurities, etc. Their number is, however, too small to give an appreciable current at ordinary field strengths. As the field strength is increased, the current due to these few electrons increases steadily, but it will not show the sudden rise observed in dielectric breakdown. For this sudden rise it is necessary that the number of electrons in an unfilled band should suddenly increase as the field strength passes a critical value. Two distinct mechanisms have been suggested for this sudden increase. Of these, the first† is a process analogous to the electrical breakdown of gases. In the absence of an external field, the few electrons in the upper band are in the lowest energy state of this band; under the action of an electric field, they are raised to higher levels. When one of these electrons reaches a sufficiently high level, it will give up energy to an electron in a lower (full) band, both electrons making a transition to a low level of the upper band. The process will then be repeated; the number of electrons in the

† Hippel, 'Z. Physik,' vol. 67, p. 707, vol. 68, p. 309 (1931); vol. 75, p. 145 (1932).

upper band will thus increase exponentially with time as long as the electric field is maintained.

When a constant electric field is present energy bands have significance only with reference to restricted regions of space, since such an electric field makes each energy band (defined for no external field) of a lattice of infinite extent degenerate with every other energy band, *cf.* fig. 1. Thus in the presence of a constant electric field an electron may pass from one "energy band" into another which, if the external field were absent, would lie above the first band. This second process of excitation is analogous† to the auto-ionization of free atoms by large electric fields. In a gas the effect of auto-ionization is of no practical importance, being masked by the first mechanism of dielectric breakdown. In a solid the analogous process is of more importance than in gases, both because the mean free path of the electrons is shorter than in gases, and because, as we shall show, the electric field can more readily *excite* an electron in a lattice than *ionize* an isolated atom.

The purpose of this paper is to examine this second method of dielectric breakdown. In § 2 the Bloch model of a metal is used to calculate the rate at which electrons escape from the lower into the upper energy bands under the influence of a constant electric field. In § 3 we show that both the magnitude of the breakdown field and the suddenness of the breakdown are obtained by this theory, assuming only the second mechanism to be operative.

## § 2. Theory.

The rate at which electrons pass from one band into another under the action of a constant electric field is independent of the thermal motion of the lattice, at least for not too high temperatures. The potential energy of an electron in the lattice will thus be considered a periodic function of the co-ordinates. Only a one-dimensional lattice will be treated, the potential energy of the electron being a function  $V(x)$  with period  $a$  of the co-ordinate  $x$  of the electron.

We shall first determine approximately the motion of an electron in a lattice under the influence of a constant electric field by the method introduced by Bloch.‡ The wave equation with a periodic potential energy has solutions of the form

$$\psi_k(x) = e^{ikx} U_k(x), \quad (1)$$

† Rosenkewitch and Sinelnikow, 'Z. Physik,' vol. 73, p. 118 (1932); Eisler, 'Z. Physik,' vol. 79, p. 266 (1932).

‡ 'Z. Physik,' vol. 52, p. 555 (1928).



where  $U_k(x)$  has the same period as  $V(x)$ . The condition that  $\psi_k$  should remain finite for all values of  $x$  imposes the restriction that  $k$  be real. But only when  $E_k$  lies in the bands of allowed energy referred to in § 1 has the wave equation solutions of the type (1) with a real  $k$ . The first energy band lies within the interval  $-\pi/a \leq k \leq \pi/a$  of  $k$ .

Bloch further showed that if, when an electric field  $F$  is present, the wave function  $\psi(x, t)$  is expanded in terms of wave functions of the lowest energy band when  $F = 0$ , in the form

$$\psi(x, t) = \int_{-\pi/a}^{\pi/a} g(k, t) \psi_k(x) dk,$$

thus neglecting all energy bands except the first and so excluding from the calculation any possibility of a transition to these bands, then

$$\frac{\partial}{\partial t} |g|^2 = -\frac{2\pi eF}{h} \frac{\partial}{\partial k} |g|^2,$$

or

$$|g|^2 = G\left(k - \frac{2\pi eF}{h} t\right),$$

where  $G$  is an arbitrary function. When we remember that  $\psi_{-\pi/a} = \psi_{\pi/a}$ , we see that  $|g|^2$  is a periodic function of  $t$ . Since the range of  $k$  is  $2\pi/a$ , and since  $|g|^2$  moves in  $k$  space with the velocity  $2\pi eF/h$ , the period of the time variation of  $|g|^2$  is  $(2\pi/a)/(2\pi eF/h) = h/eFa$ ; its frequency is thus

$$\nu = eFa/h. \quad (2)$$

We thus conclude that the velocity of the electron, *i.e.*,

$$\frac{\partial}{\partial t} \int \psi x \psi^* dx,$$

is also a periodic function of time with the frequency  $\nu$ , since the identity

$$\frac{\partial}{\partial t} \int \psi x \psi^* dx = \int v_k |g|^2 dk$$

may be readily established when  $v_k$  is the velocity of an electron in state  $k$ . Thus if we represent the electron by a wave packet confined to the first energy band, the electron moves in the direction of the field until it is reflected by the lattice, then moves in the opposite direction until it is stopped by the field, whereupon the motion is repeated.

We must now calculate the probability per unit time that the electron will make a transition to an excited band. If we take the wave function  $\psi(x, t)$  to be a wave packet, then each time the wave packet is reflected by the lattice, a small fraction of the wave packet will pass through into the upper zone, giving the probability  $p$  that the electron escapes into the upper zone (cf. fig. 1). The probability,  $\gamma$ , per unit time is obtained by multiplying  $p$  by the

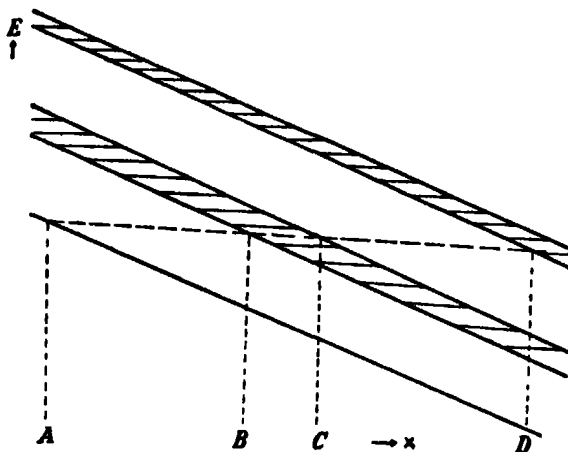


FIG. 1.—“Potential barrier” diagram. The shaded regions represent zones of forbidden energy in the presence of an electric field.

number of times per second that the wave packet is reflected by the lattice, namely,  $\nu$ , so that

$$\gamma = \nu p.$$

We have now to calculate  $p$ . This is most simply done by taking for  $\psi(x, t)$  a wave function *periodic* (in time), not a wave packet. The wave equation for an electron with energy  $E$  when an external electric field  $F$  is present is

$$\left\{ \frac{\hbar^2}{8\pi^2 m} \frac{\partial^2}{\partial x^2} - V(x) + (E - eFx) \right\} \psi(E : x) = 0.$$

An approximate solution of this equation may at once be obtained if we observe that the term  $eFx$  changes only slightly in the distance of a single lattice constant. Such a solution is

$$\psi(E : x) = e^{i \int K dx} U_k(x),$$

where  $K$  is the same function of  $(E - eFx)$  as it is of  $E$  alone when  $F$  is zero, and  $U_k$  is identical to the  $U_k$  of (1) with  $k = K$ . With a given  $E$  our  $K$  will now be real *only for certain values* of  $x$ , as shown in fig. 1. In the regions

where  $K$  is complex,  $\psi(E : x)$  will decrease, or increase, exponentially. On either side of such a region  $\psi(E : x)$  is nearly identical with the wave functions in the absence of a field which belong to two adjacent energy bands. Thus in fig. 1,  $\psi(E : x)$  is similar to a wave function belonging to the first energy band in the interval AB, while in the interval CD it is similar to a wave function belonging to the second band. If we take the wave function  $\psi$  which *decreases* in the range BC, then we know that the transition probability  $p$  is given by†

$$p = |\psi(x_C)/\psi(x_B)|^2.$$

If we write, for values of  $x$  in the interval BC,

$$K = \xi(x) + i\eta(x) \quad (4)$$

then we have

$$|\psi(x_C)/\psi(x_B)| = e^{-\int_B^{x_C} \eta(x) dx} \quad (4a)$$

and hence, by (2) and (3)

$$\gamma = \frac{eFa}{h} e^{-2\int_B^C \eta dx}.$$

In order to obtain an explicit expression for  $K(x)$ , we write our wave equation in the standard form of Hill's equation,

$$\frac{d^2\psi}{ds^2} + (0_0 + 2\sum_{n=1}^{\infty} 0_n \cos 2ns) \psi = 0.$$

Here

$$s = \pi x/a,$$

$$0_0 = 8ma^2(E - eFz)/h^2,$$

and  $\theta_m$  is  $4ma^2/h^2$  times the coefficient of  $\cos 2\pi nx/a$  in the Fourier expansion of  $V(x)$ . When  $\theta_1$  is small in comparison with unity (with  $a = 3 \times 10^{-8}$  cm, this condition means that the first Fourier coefficient must be small in comparison with 4 electron volts), and when the higher coefficients are still smaller,  $K$  is determined as a function only of  $\theta_0$  and  $\theta_1$  by the equation‡

$$\sin^2(aK/2) = \sin^2(\frac{1}{2}\pi\theta_0^{\frac{1}{2}}) \left\{ 1 + \frac{\pi\theta_1^2}{4\theta_0^{\frac{1}{2}}(1-\theta_0)} \cot(\pi\theta_0^{\frac{1}{2}}/2) \right\} + O(\theta_1^4). \quad (5)$$

In the first energy band in which  $K$  is complex,  $\theta_0$  is seen from (5) to differ

† Gamow, "Atomic Nuclei and Radioactivity," chap. 2, Oxford (1931); v. Laue, 'Z. Physik,' vol. 52, p. 726 (1928). Houtermans, 'Ergbn. exakt. Naturw.,' vol. 9, p. 133 (1930).

‡ Hill, 'Acta Math.,' vol. 8, p. 27 (1886); M. J. O. Strutt, "Lame'sche-Mathieusche Funktionen," p. 25, Berlin (1932). In the latter reference  $2\theta_1$  is erroneously used in place of  $\theta_1$ . This mistake is easy to make, since  $\theta_1$  is defined in Hill's memoir in two ways differing by a factor of two.

only slightly from unity, and  $K$  is seen to differ only slightly from  $\pi\theta_0^{\frac{1}{2}}/2$ , i.e., only slightly from  $\pi/a$ . Hence if we set

$$\left. \begin{aligned} \theta_0 &= 1 + \alpha \\ K &= \pi/a + \beta \end{aligned} \right\} \quad (6)$$

$\alpha$  and  $\beta$ , together with  $\theta_1$ , may be regarded as small quantities. Substituting (6) in (5), and neglecting powers of  $\alpha$ ,  $\beta$ , and  $\theta_1$  greater than the second, we obtain

$$K = \pi/a \pm i(\pi/a)(\theta_1^2 - \alpha^2)^{\frac{1}{2}}.$$

If we now choose the origin  $x = 0$  to lie midway between B and C, we have, using (6),

$$K = \frac{\pi}{a} \left[ 1 \pm i \frac{8ma^2}{\hbar^2} \left\{ V_0^2 - (eFx)^2 \right\}^{\frac{1}{2}} \right]. \quad (7)$$

The points B and C are now given by  $x = -V_0/eE$  and  $x = V_0/eF$ , respectively, so that the energy gap  $\epsilon$  between the two bands of allowed energy is  $2V_0$ . Performing the integration of (4a), we obtain for the rate  $\gamma$  with which an electron passes from the region AB into CD

$$\gamma = \frac{eFa}{\hbar} \exp - \left\{ \frac{\pi^2 ma \epsilon^2}{\hbar^2 |eF|} \right\}. \quad (8)$$

### § 3. Discussion.

We have obtained in (8) the rate with which an electron escapes from one energy band into the next energy band. In the derivation of (8) we assumed the system to be one-dimensional, and the first Fourier coefficient of the potential energy of an electron in the lattice to be sufficiently small (considerably less than 4 electron volts when  $a = 3 \times 10^{-8}$  cm). It is to be expected that the results obtained for a one-dimensional lattice are true in general for a three-dimensional lattice. Although the Fourier coefficients of most crystals are larger than allowed in the above derivation, it is nevertheless to be expected that the general dependence, given by (8), of the transition rate upon the energy gap  $\epsilon$  will still be valid.

When we take the reasonable magnitudes  $a = 3 \times 10^{-8}$  cm,  $\epsilon = 2$  electron volts, and let  $F^*$  be the electric field in volts  $\text{cm}^{-1}$ , (8) becomes

$$\gamma = 10^7 F^* 10^{-2 \times 10^7/F^*}.$$

The dielectric breakdown is thus not to be expected until  $F^*$  reaches the order of magnitude of  $10^6$ , as is observed. Further, the breakdown will occur

suddenly as  $F^*$  is increased,  $\gamma$  being increased by a factor of 100 (in our example) when  $F^*$  changes from  $1.0 \times 10^6$  to  $1.1 \times 10^6$ .

It is to be particularly stressed that for field strengths obtainable in metals, equation (8) shows the transition between different energy bands to be altogether negligible. Thus for  $a = 3 \times 10^{-8}$  cm, and  $F^* = 1$ , an electron will pass an energy gap as small as  $2 \times 10^{-3}$  electron volts only after a period of years.

#### § 4. *Summary.*

A theory of the electrical breakdown of solid dielectrics is developed. The breakdown is assumed to be due to the direct excitation of electrons by the electric field. Agreement with experiment is obtained both for the magnitude of the breakdown field and for the suddenness with which the breakdown occurs as the electric field is increased.

---

### *Some Peculiarities in the Physical Properties of Iron-Aluminium Alloys.*

By C. SYKES, Ph.D., and H. EVANS, Assoc. Met.

(Communicated by W. L. Bragg, F.R.S.—Received December 28, 1933.)

The thermal equilibrium diagram of alloys consisting of iron and aluminium has been the subject of several investigations from which it appears that solid iron forms a solid solution with aluminium up to 30% by weight of aluminium.\* Whilst carrying out an examination of the change in physical properties of iron with increasing aluminium content, from 0% to 17%, it was noticed that certain physical properties, notably thermal expansion and electrical resistivity, exhibited certain peculiarities not to be expected from a range of solid solutions. It was thought worth while to examine these effects in greater detail and the results are recorded in this paper.

#### EXPERIMENTAL.

##### *Production of Alloys.*

The specimens were made from Swedish iron or conductor rail and high grade aluminium, and melted in a high frequency furnace in air. They were

\* Gwyer and Philips, 'J. Inst. Met.,' vol. 38, p. 29 (1927).

cast into chill moulds and forged or swaged to suitable shape. The major impurities were silicon 0.3% and carbon 0.06%.

### *Thermal Expansion.*

A series of specimens were made containing from 0–16% aluminium, they were 2 inches in length and  $\frac{3}{8}$  inch in diameter. The extension was measured at temperatures up to 1000° C., and when this was plotted against temperature it was found that certain of the alloys, namely, those containing an aluminium content in excess of 12%, gave a curve which showed a change in direction at temperatures in the range 500° C.–600° C. For alloys containing 12.2% and 13.3% the change occurred at about 560° C., and for the 15.5% alloy it occurred at 500° C.

From these curves the mean coefficient of expansion was determined up to various temperatures and has been plotted against aluminium content, fig. 1. It is clear from this figure that alloys containing from 12% Al upwards exhibit some peculiarity in the range 500°–600° C.

It was natural to suppose that some form of transformation was taking place in these alloys giving rise possibly to a small volume change. Further experiments were then carried out using a much larger specimen 10 inches long in a more sensitive apparatus and the expansion curve was measured using very slow rates of cooling in the temperature range 700° C.–500° C. Fig. 2 shows the results obtained from which it appears that no discontinuity in volume takes place, merely a change in the rate of expansion.

### *Resistivity Measurements.*

A series of specimens were examined containing up to 16% aluminium; beyond this the alloys are very brittle and cannot be worked either hot or cold. Up to 11% the values of specific resistance varied practically linearly with aluminium content and no difficulty was experienced in repeating results with different samples. Beyond this no concordant results were obtained for some time until it was noticed that the resistivity depended on the previous thermal history of the specimen, i.e., the resistivity of a sample when quickly cooled was not the same as when it was slowly cooled from red heat.

In Table I is given the resistivity of various alloys after different heat treatments. It is seen that the alloys from 11% downwards have the same resistivity independent of the heat-treatment, whilst alloys richer in aluminium show a marked variation with cooling rate. The effect is reversible.

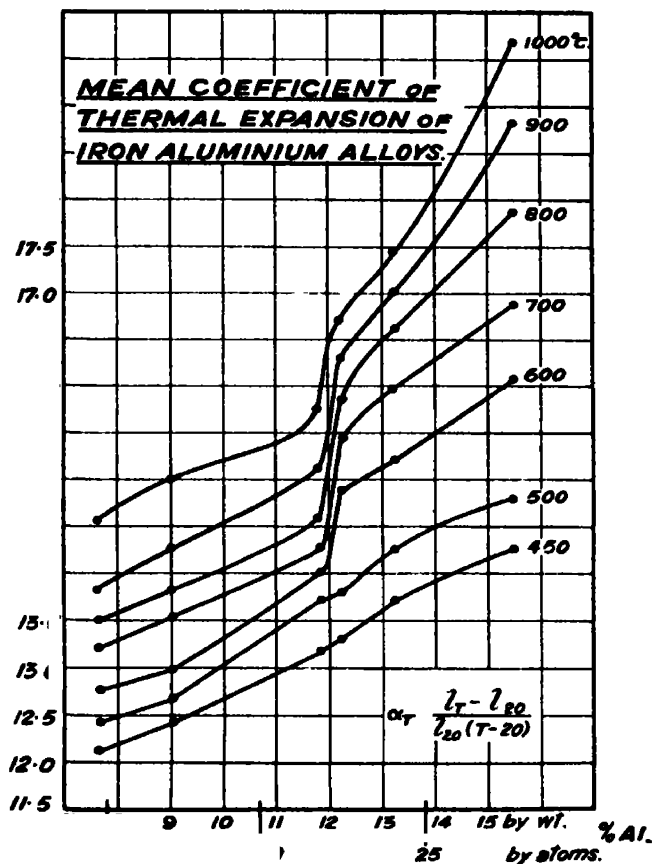


FIG. 1.

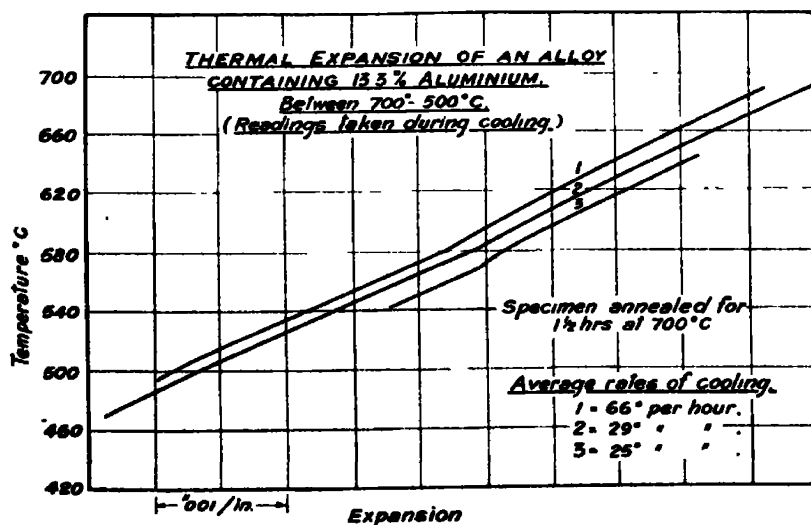


FIG. 2.

Table I.

Aluminium content % by weight.	Resistivity after treatment, ohms $\times 10^{-8}$ per cm. cu.			
	1.	2.	3.	4.
7.3	75	75	75	75
10.3	92	92	93	92
11.2	104	104	105	105
13.3	128	88	128	91
15.4	139	98	141	100

Treatment 1.—Annealed at 700° C. for 20 minutes and then quenched in water.

Treatment 2.—Annealed at 1000° C. for 1 hour and cooled at 30° C. per hour to room temperature.

Treatment 3.—Repeat of treatment 1 after treatment 2.

Treatment 4.—Repeat of treatment 2 after treatment 3.

Table II lists resistivities of other alloys which were annealed for long periods in the range 450° C.–550° C. in order to determine whether any further modification in resistivity took place. In all samples which had previously been

Table II.

Aluminium content by weight :	11.2.	11.8.	12.2.	13.1.	13.3.	14.4.	15.5.	16.5.
Resistivity after treatment 5 . . . . .	105	104	98	96	—	97	—	140
Resistivity after treatment 6	105	106	98	91	90	89	102	140
Resistivity after treatment 7	105	106	119	129	125	135	147	157

Treatment 5.—Heated to 720° C. and heated for 1½ hours, cooled to 550° C. in 10 hours, annealed for 10 hours, cooled to 500° C. in one hour, and then annealed for 24 hours, cooled to room temperature in 24 hours.

Treatment 6.—Annealed at 1000° C. for 30 minutes, cooled to 500° C. at 30° C. per hour, annealed for 36 hours, cooled to 450° C., annealed for 72 hours, cooled to room temperature in 16 hours.

Treatment 7.—Annealed at 1000° C. for 30 minutes, cooled to 700° C., annealed for 1½ hours and then quenched in water.

quenched the effects of the longer annealing periods were negligible, *e.g.*, compare the resistivity of annealed samples of similar composition in Tables I and II. The two samples containing 13.1% Al and 14.4% Al respectively had not been quenched prior to heat treatments 5 and 6; after treatment 7 they were again heat treated by cooling at 30° C. per hour from 700° C. The results were identical with those obtained from treatment 6. After a further quenching they were submitted to another prolonged heat treatment similar to treatment 6 and the results were again in agreement with the previous results.

The accuracy of measurement of resistance in these experiments was of the order of  $\pm 1\%$ . Bearing this in mind consider the resistivity of the samples



containing 13.3% and 15.4% Al in Table I. The agreement is not good. This lack of complete reversibility is not due to oxidation of the specimens as such alloys resist oxidation very well indeed, *e.g.*, alloys which had had all the various treatments were only affected to the extent of a very fine red film of oxide. On the other hand, the alloys are very brittle, standard izod impact tests give values of the order of 1 ft. lb., consequently there is considerable risk of cracking the specimens during quenching. This difficulty is accentuated the higher the aluminium content, and renders investigation of alloys containing more than 15% aluminium almost impossible by this means.

Quenching in oil from 1000° C. instead of from 700° C. in water did not produce satisfactory results, the values of resistivity being much lower than the values obtained by water quenching.

Fig. 3 shows the variation of electrical resistivity with aluminium content, the alloys being in two conditions: (1) as quenched from 700° C. in water, and (2) after cooling at 30° C. per hour. Depending on the rate of cooling any value between the limits given in fig. 3 are possible.

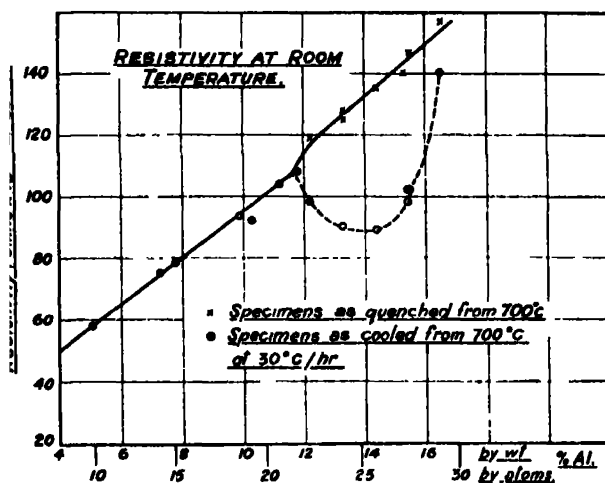


FIG. 3.

Various experiments were then carried out to determine if this change in resistivity took place at any definite temperature. Quenching at temperatures from 1250° C. to 600° C. did not affect the resistivity as quenched, *e.g.*,

14.4%	alloy	quenched from	1250° C.,	resistivity	134.
14.4%	"	700° C.,	"	135.	
13.2%	"	700° C.,	"	129.	
13.2%	"	600° C.,	"	129.	

Variations in resistivity were noticed when the quenching temperature was lowered to 500° C.

Table III shows the effect of lowering the quenching temperature, the treatments were carried out in the order given on a sample containing 13·3% Al.

The specimen was soaked at 700° C. for 30 minutes before the subsequent treatment, viz., cooling at 30° C. per hour to the quenching temperature, was commenced.

From the results in Table III it appears that the change in resistivity takes place over a range in temperature, and, in fact, whilst it commenced between 600° C. and 500° C. it is not complete at 400° C.

Table III.

Treatment.	Resistivity.
Specimen cooled from 700° C. at 30° C. per hour	89
„ quenched from 700° C.	129
„ „ „ 600° C.	128
„ „ „ 500° C.	115
„ „ „ 400° C.	102
„ „ „ 200° C.	89
„ cooled from 700° C. at 30° C. per hour	88

In order to check this point with certainty a series of alloys were investigated at temperatures up to 900° C., the resistance being measured with the specimen hot. The results are shown in fig. 4. Here the proportionate increase in resistance is plotted against temperature.

Apart from the abnormal increase in resistivity at present under discussion there is the possibility that the alloys possess an appreciable temperature coefficient (for example, an alloy containing 10% aluminium has a temperature coefficient of 0·0008 between 20° C. and 600° C.), consequently the precise significance of the curves is difficult to assess. The main point of importance for the present is that the curves do not show a discontinuous change in resistivity at any one temperature. The cooling time for the specimens in these experiments was 100 hours. At each temperature at which readings were taken the specimens were maintained at constant temperature until no difference in resistance could be measured over a period of several hours. The agreement obtained is shown in the following examples. The different specimens are numbered 1, 2, 3, 4, 5. The analyses were: 11·4%, 13·6%, 14·5%, 15·9%, 17·4% aluminium respectively.

Table IV.

Temperature.	Time.	Factor proportional to resistance of specimen.					Time interval.
		1.	2.	3.	4.	5.	
719	5.30 p.m.	11.90	10.87	14.09	12.60	12.81	15 hours.
718	8.45 a.m.	11.91	10.87	14.09	12.64	12.82	
504	5.45 p.m.	10.23	10.02	13.79	12.69	12.75	15 hours.
504	9.00 a.m.	10.20	10.09	13.82	12.76	12.76	

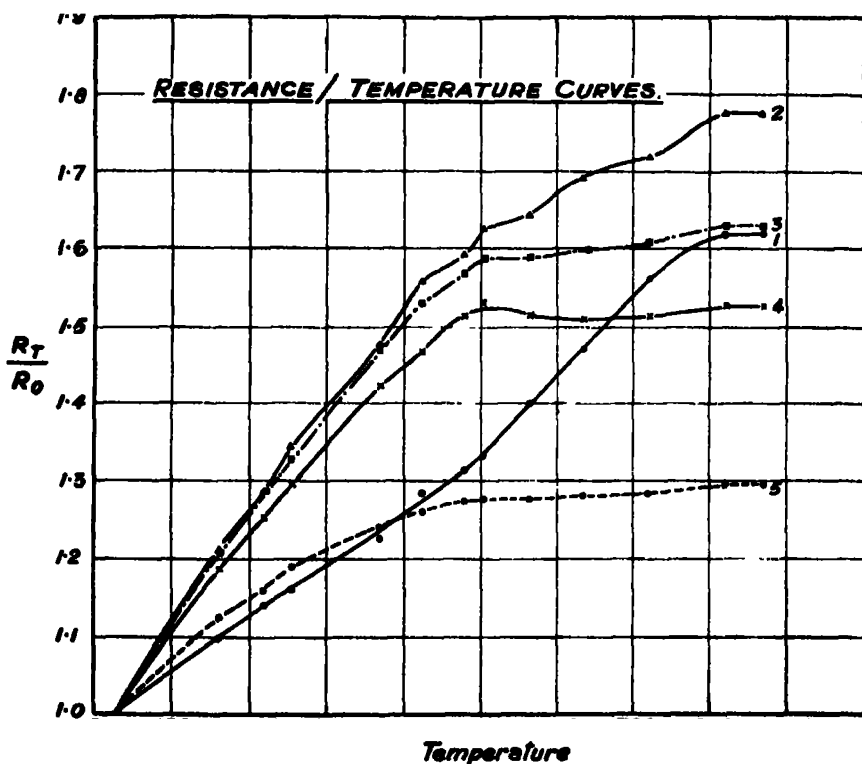


FIG. 4.

From the results at 719° C. it is clear that oxidation effects, if any, are extremely small. Previous results have indicated that the transformation has commenced at 504° C., see Table III, in the alloy containing 13.3% aluminium. Specimen 2 has practically the same composition as this alloy and yet there is no sign of any transformation completing itself during the 15 hours the specimen was held at this temperature. In fact, further annealing

at any one temperature did not produce any change of resistance after the specimen had been cooled at a rate of 30° C. per hour.

### CONCLUSIONS.

With the exception of the resistivity changes the effects noted are small. In addition we have investigated the variation of hardness and tensile strength in this range of alloys with particular reference to heat-treatment. Such variations which we have noticed are within the limits of experimental error and it cannot be said that the transformation affects these properties appreciably.

The resistivity experiments are conclusive, however, in showing that alloys containing 12% aluminium upwards to 16% aluminium can be modified by heat-treatment. The effect appears to be a maximum between 13% and 14% aluminium by weight, *i.e.*, in the region of the alloy of composition  $\text{Fe}_3\text{Al}$ .

The type of transformation usually met with in alloys is one which takes place at a definite temperature and provided the alloy is cooled reasonably slowly through this temperature or annealed at a temperature slightly below the transformation temperature, the change is substantially completed. Certain solid transformations require very long periods in order to become complete. If such a transformation involving very slow cooling for completion were occurring in this system of alloys, we should naturally expect to have found evidence of a gradual change in properties during the various annealing experiments carried out to settle this particular point, and certainly in the experiments connected with measurement of resistance with the specimens hot. For example, consider the alloy containing 13.3% aluminium. Treatment (1) and treatment (6) applied to this alloy produce similar effects on the resistance. However (see Table III), it is known that if treatment (1) is stopped at 500° C. and the alloy quenched the resistivity is  $115 \times 10^{-6}$  ohms, indicating that the transformation has started but is not complete. Treatment (6) is similar to treatment (1) with the addition of 36 hours annealing at 500° C. and 72 hours at 450° C. The fact that both treatments produce identical results suggest two possibilities. In the first place it can be argued that the transformation takes place at a definite temperature under very slow rates of cooling, *i.e.*, slower than 30° C. per hour, consequently in treatment (1) undercooling occurs and the change takes place over a range of temperature, but is substantially completed at room temperature. In treatment (6) this argument indicates that the transformation would be finished at a higher temperature

than in treatment (1), but that the final result would be the same. If this explanation is correct, the resistivity of a specimen cooled to 500° C. and then held at this temperature should change with time, this is not so. We are of the opinion, therefore, that the experimental evidence indicates definitely the second possibility, namely, that the transformation takes place over a range of temperatures, and that at any one temperature in that range it proceeds to the corresponding partial completion quickly; for instance, a cooling rate of 30° C. per hour is sufficient to ensure that it takes place. The range of temperature in a 13% alloy stretches from 500° C. to below 400° C.

Bradley and Jay,\* by means of X-ray analysis, have made a study of this system of alloys. They have shown that under slow cooling conditions, about 30° C. per hour, rearrangement takes place amongst the atoms and a symmetrically arranged solid solution is obtained. The type of lattice, i.e., body-centred cubic, was the same throughout.

According to them the alloys can be divided into two groups. In the first group containing from 10–13·9% aluminium, the atoms in the alloy quenched from 600° C. are arranged at random, whereas under conditions of slow cooling the alloys exhibit a superlattice indicating that the atoms have emerged from the random arrangement into a more or less ordered arrangement. This becomes complete at 13·9% aluminium corresponding to the composition  $\text{Fe}_3\text{Al}$ . In this alloy the aluminium atoms are situated in definite positions in the iron lattice. In the second group containing from 14% upwards, the alloys as quenched exhibit a more or less ordered structure similar to that obtained for the compound  $\text{FeAl}$ , while the structure of the annealed alloys gradually changes from the  $\text{Fe}_3\text{Al}$  type to the  $\text{FeAl}$  type.

The general change of resistivity over the range of alloys investigated agrees very well with that expected if we assume that the alloy with atoms in the ordered arrangement has a lower resistivity than the same alloy having a random arrangement of atoms. In fact, the general agreement is so good that it seems reasonable to conclude that the rearrangement of the atoms is the cause of the change in resistivity.

If this is so, then the atomic rearrangement must also take place over a range of temperature attaining equilibrium at any one temperature in a reasonably short time. At any particular temperature in this range a corresponding definite degree of order can be attained which prolonged annealing at this temperature will not modify. Any further change to a more ordered state can then be brought about only by reducing the temperature.

\* 'Proc. Roy. Soc.,' A, vol. 136, p. 210 (1932).

This conclusion is confirmed by the following point :—

According to Bradley and Jay (*loc. cit.*) the lattice spacing of the annealed material is different from that of the quenched material for a 13·3% alloy by about 0·1%. This would correspond to a change in length on a 10-inch specimen of 0·01 inch. If this took place at a particular temperature, it is hardly likely to have been missed in the experiments already described, and consequently we must assume that the change of lattice spacing consequent on the rearrangement of the atoms takes place over a definite range of temperature.

Certain other alloy systems, notably copper-gold, copper-palladium, are similar to the iron-aluminium system in that the resistivity of the alloys can be modified by heat treatment. The effect is most marked at compositions corresponding to simple atomic proportions, i.e.,  $\text{Cu}_3\text{Au}$ ,  $\text{Cu}_3\text{Pd}$ ,  $\text{CuAu}_3$ , and  $\text{CuPd}$ . In the alloys  $\text{CuAu}$  and  $\text{CuPd}$  not only does a rearrangement of atoms in the lattice occur, but the type of lattice also changes; consequently such transformations are not exactly comparable with those of alloys of the type  $\text{Cu}_3\text{Au}$ ,  $\text{Cu}_3\text{Pd}$ ,  $\text{Fe}_3\text{Al}$  where the type of lattice is not changed. The copper-gold alloys have been the subject of much experimental work.

Horton and Payne\* state that the transformation in the  $\text{Cu}_3\text{Au}$  alloys takes place over a range of temperature of about 15°–20° C. and that whilst the upper limits are sharp the lower limits are vague irrespective of whether heating or cooling curves are concerned. This is interesting since it suggests that the change from random arrangement into a state of order sets in rapidly. The cooling rates used in this work were of the order of 5° C. per hour, so that the rate at which the transformation completes itself is very slow indeed, at least 3 hours seems essential.

In many ways the  $\text{Cu}_3\text{Au}$  transformation appears to resemble that described for  $\text{Fe}_3\text{Al}$ , and it would be very interesting to know if the range of temperature over which the change takes place is really a function of the cooling rate.

The authors gratefully acknowledge the assistance of Captain J. W. Bampfyde, of the British & Dominions Feralloy Co., Ltd., in connection with measurements of expansion and analyses of samples. Their thanks are also due to Mr. A. P. M. Fleming, C.B.E., Director of Research and Education Departments, Metropolitan-Vickers Electrical Co., Ltd., and Mr. C. P. Sandberg, O.B.E., of the British and Dominions Feralloy Co., for permission to publish this paper.

The authors are indebted to Professor W. L. Bragg for his kind interest and encouragement.

\* 'J. Inst. Met.,' vol. 48, p. 457 (1931).

SUMMARY.

An account is given of measurements of the resistivity of alloys of iron and aluminium containing from 11–16% aluminium by weight. It is shown that the resistivity of such alloys at room temperature depends on the rate of cooling of the specimens from a temperature of the order of 600° C.

Alloys in this range consist of a single solid solution at all temperatures concerned. It is concluded, therefore, that rearrangement of atoms takes place in the alloys under slow cooling conditions, and the more regular arrangement so produced leads to a decrease in resistance.

Experimental results are described suggesting that the rearrangement of atoms in the space-lattice takes place over a considerable range of temperatures even under conditions of very slow cooling.

---

*The Dielectric Constants of Liquid Mixtures of Phenol—Water, Phenol—m-Cresol, Phenol—Aniline, and Phenol—p-Toluidine*

By OWEN RHYS HOWELL and WILLIS JACKSON, The College of Technology, Manchester

(Communicated by T. M. Lowry, F.R.S.—Received January 3, 1934.)

*Introduction*

The densities,\* viscosities,† and electrical conductivities‡ of solutions in the system phenol–water have been measured over the whole range of miscibility at 20°, 30°, 40°, 50°, 60°, and 70° C. It has been shown that at each temperature the variation of each of these properties with composition of the mixture is quite continuous and gives no indication of the presence in solution of phenol-hemihydrate.

On the contrary, the constants in the equations expressing the variation of each of these properties with temperature, when plotted against the composition, exhibit marked inflexions at the concentration corresponding to

\* Howell, 'Proc. Roy. Soc.,' A, vol. 137, p. 418 (1932).

† Howell, 'Trans. Faraday Soc.,' vol. 28, p. 912 (1932).

‡ Howell and Handford, *ibid.*, vol. 29, p. 640 (1933).

equimolecular proportion of the two constituents. This is attributed to inter-ionization of the two substances through the transfer of a hydrogen ion from a molecule of one component to a molecule of the other.

The electrical conductivities of the systems phenol—*m*-cresol, phenol—*aniline*, and phenol—*p*-toluidine\* have also been examined at 50°C. The conductivity curves, again, give no indication of the presence of the compounds shown on the corresponding freezing-point curves but, on the contrary, suggest that inter-ionization may occur in other molecular proportions.

The conductivity—composition curves of these binary systems exhibit very pronounced maxima (except phenol—*m*-cresol, where the effect is not so marked) and it was therefore deemed of interest to examine the dielectric constants to ascertain how far the variation in electrical conductivity is dependent on the simultaneous change in the dielectric constant. It has been found that there is no parallelism between the change in the two properties, but the dielectric constant—composition curves exhibit very marked discontinuities, which are discussed.

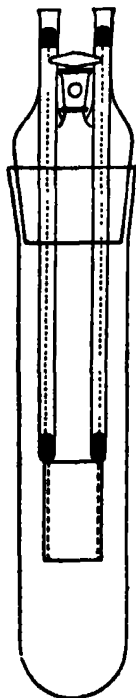


FIG. 1

### *Experimental*

*Cell*—The cell used for all the measurements is shown in fig. 1. It consisted of a tube, 5" long and  $1\frac{1}{4}$ " diameter, of borosilicate glass, fitted with a ground-glass stopper carrying two electrodes (the plates of the test condenser) at the ends of two leading tubes. The electrodes were co-axial cylindrical platinum plates, each 2.5 cm long and respectively of 1.5 cm and 1.3 cm diameter,

held rigidly by four small glass beads fused on at both top and bottom. Connection was made with the electrodes by two stout platinum wires fused into the leading tubes. The ends within the tubes were covered with mercury into which dipped two platinum wires which ran up the tubes and terminated in two mercury cups at the top. Electrical connection between these and two other mercury cups terminating the other part of the circuit was established by means of a connecting link consisting of two stout copper rods held rigidly in ebonite cross-pieces, thus ensuring that there was no variation in the capacity of the external circuit as a result of the making and breaking of connection with the cell necessary for each determination.

\* Howell and Robinson, 'J. Chem. Soc.,' p. 1032 (1933).



*Solutions*—The materials were all purified in the manner already described.\* The solutions were made up by weight directly in the tube used for the determination. There were six such tubes, the stopper with the electrodes being interchangeable for all, so that preparation and measurement might proceed simultaneously.

*Temperature*—All measurements were made in a large thermostat filled with water covered with a layer of oil and heated by gas from a pressure-governed supply, the temperature being constant to within  $0.01^{\circ}\text{C}$ . The phenol—*m*-cresol, phenol—*aniline*, and phenol—*p*-toluidine systems were examined at  $50^{\circ}\text{C}$  and the phenol—water system at  $70^{\circ}\text{C}$ , since it was necessary in this instance to work above the critical solution temperature,  $66.4^{\circ}\text{C}$ .

The tube containing the solution under examination was closed with a plain stopper and immersed up to the base of the stopper in the thermostat for half an hour. The stopper carrying the electrodes was placed in a similar empty tube, which was similarly immersed to enable the electrodes also to acquire the requisite temperature. This stopper was then transferred to the tube containing the solution and a reading was taken immediately. Other readings were taken at intervals to ensure that the temperature of the thermostat had been attained.

After the measurement, the electrodes were thoroughly washed with acetone, with hot water and again with acetone, dried in a stream of air and returned to the empty tube in the thermostat. The tubes were similarly cleaned before being used for the preparation of further solutions.

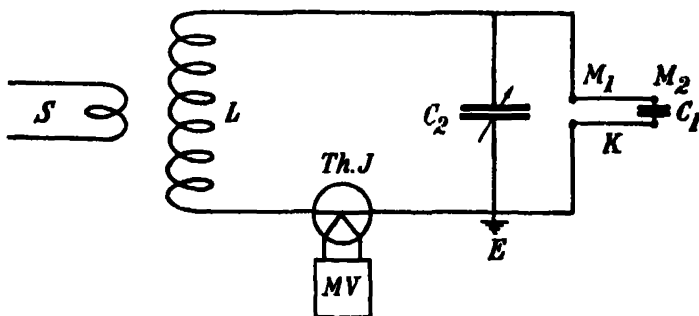


FIG. 2

*Electrical Circuit*—Measurements were made at a frequency of  $10^6$  cycles per second using a resonance method. The circuit is shown in fig. 2. The test condenser *C<sub>1</sub>* is in parallel with a calibrated variable air condenser *C<sub>2</sub>* having

\* Howell and Handford, Howell and Robinson, *loc. cit.*

vernier adjustment and reading to  $0.1 \mu\text{f}$ , connection being made by the removable link K dipping into the pairs of mercury cups  $M_1$  and  $M_2$ .

The two condensers,  $C_1$  and  $C_2$ , form with the coil L a series resonant circuit, the current flow in which was measured on a vacuo-thermal junction ThJ and its associated millivoltmeter MV. Voltage of the desired frequency was injected into L from a source of continuous oscillations S, the frequency of which could be adjusted and maintained accurately constant by comparison with a heterodyne wavemeter. The output of the valve oscillator employed was sufficient to permit of very weak coupling to the measuring circuit, except for the most conducting of the phenol—water samples, so that, with this exception, the possibility of feed-back on to the oscillator did not arise. This is important since the accuracy of the dielectric constant determinations depends on the maintenance of constant oscillator output and frequency as the resonance point of the measuring circuit is traversed.

The capacitance added to that of the condenser  $C_2$  on connection of the test condenser  $C_1$  with air as dielectric consists of two parts

- (1) a part  $C_1$  due to the leads and to those portions of the plates which are embedded in the glass beads. This part may be regarded as forming permanently a constant addition to  $C_2$ ;
- (2) a part  $C_1$  between the plates, which on replacement of the air by a liquid of dielectric constant  $\epsilon$ , becomes  $\epsilon C_1$ .

If therefore the values of the variable condenser  $C_2$  required to give resonance of the measuring circuit at the testing frequency are  $C'_2$  and  $C''_2$  for the air and liquid dielectrics respectively, then

$$C'_2 - C''_2 = (\epsilon - 1) C''_1.$$

The measurement in the test condenser  $C_1$  of a liquid of known dielectric constant enables  $C''_1$ , the cell constant, to be calculated and, from a knowledge of this, unknown dielectric constants can then be determined.

*Cell Constant*—Comparatively few values are available for dielectric constants at the working temperatures of this investigation, so that the choice of liquid for determination of the cell constant is limited. It was considered best to use aniline, for which the following values are recorded:  $\epsilon = 6.59$  (interpolated),\*  $6.43$  (interpolated)† and  $6.295$ ‡ at  $50^\circ \text{C}$  and  $6.20$ § at  $58^\circ \text{C}$ . Accepting  $\epsilon = 6.30$  as the most probable value, the cell constant  $C''_1$  is  $12.30$ .

\* Ratz, 'Z. phys. Chem.,' vol. 19, p. 94 (1896).

† Jezewski, 'J. Phys. Radium,' vol. 3, p. 293 (1922).

‡ Estermann, 'Z. phys. Chem.,' B, vol. 1, p. 134 (1928).

§ Kerr, 'Phil. Mag.,' vol. 3, p. 330 (1927).

Calculation shows that the cell constant should not vary appreciably with temperature and this value was therefore used for all the measurements both at 50° and 70° C.

All the dielectric constants now recorded, therefore, are based on this accepted value for aniline at 50° C, and, as absolute values, are dependent on its accuracy. This does not, however, detract from the relative accuracy of the values among themselves, nor from the significance of the results and the conclusions drawn from them. The experimental error in the values of  $\epsilon$  is about  $\pm 0.01$ , but may be somewhat greater for the most highly conducting solutions.

That the values now given are not far removed from the true values is seen by comparing those for phenol and *p*-toluidine with the values found by other observers :—

Phenol : previously recorded,  $\epsilon = 9.90^*$  at 58° C. Using the differences in the readings now found at 50° C and 70° C to apply the temperature correction, the value is  $\epsilon = 10.36$  at 50° C, as against 10.28 now recorded.

*p*-Toluidine : previously recorded,  $\epsilon = 4.88^*$  at 58° C and  $\epsilon = 5.40^\dagger$  at 44° C. Interpolating,  $\epsilon = 5.10$  at 50° C as against 5.07 now recorded.

*Conductivity Correction*—The equation given above is based on the assumption that the liquid dielectric is devoid of conductivity. Since the liquids investigated are known to possess a marked and varied conductivity, it is necessary to ascertain whether a correction must be applied. When filled with liquid, the test condenser  $C_1$  may be regarded, together with  $C_2$ , as shunted by a leakage resistance  $R$ , as shown in fig. 3*a*. This combination can be reduced to the equivalent series resistance circuit of fig. 3*b* in which  $r$  is the series resistance added to the circuit by the conductivity of the test condenser and  $C_s$  is the equivalent series capacitance. For the two circuits to be identical at a frequency of  $w/2\pi$  cycles per second, the real and imaginary parts of their symbolic impedances must be equal respectively.

For the circuit of fig. 3*a*,

$$Z = jwL - \frac{R \left[ \frac{j}{w(C_1 + C_2)} \right]}{R - \frac{j}{w(C_1 + C_2)}} \\ = \frac{R}{1 + R^2(C_1 + C_2)^2 w^2} + j \left\{ wL - \frac{R^2(C_1 + C_2) w}{1 + R^2(C_1 + C_2)^2 w^2} \right\}.$$

\* Kerr, *loc. cit.*

† Cauwood and Turner, 'J. Chem. Soc.,' vol. 107, p. 276 (1915).

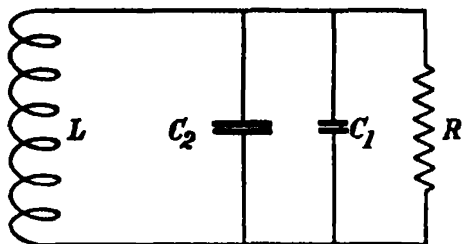
For the circuit of fig. 3b,

$$Z = r + j \left\{ \omega L - \frac{1}{\omega C_s} \right\}.$$

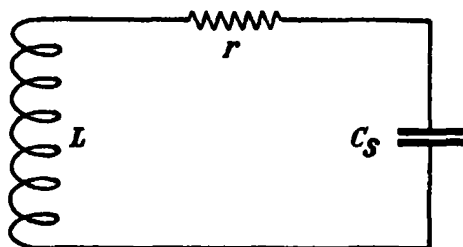
Hence equating the real and the imaginary parts,

$$r = \frac{R}{1 + R^2 (C_1 + C_2)^2 \omega^2}$$

$$C_s = \frac{1 + R^2 (C_1 + C_2)^2 \omega^2}{R^2 (C_1 + C_2) \omega^2} = (C_1 + C_2) + \frac{1}{R^2 (C_1 + C_2) \omega^2}.$$



a.



b.

FIG. 3

When \$R\$ is infinite, corresponding to no conductivity, \$C\_s\$ is equal to \$C\_1 + C\_2\$. Since the adjustment of the circuit to resonance by variation of \$C\_2\$ does not depend directly on \$C\_2\$ but on the value of the equivalent series capacitance \$C\_s\$, the value of \$C\_2\$ required to give resonance will be less for a conducting liquid than for the same liquid in the absence of conductivity by an amount equal to  $\frac{1}{R^2 (C_1 + C_2) \omega^2}$ . This quantity must therefore be subtracted from the observed value of \$(\epsilon - 1) C''\_1\$.

Since the conductivities of solutions in all the binary systems examined are already available over the whole range of composition,\* this correction factor

\* Howell and Handford, Howell and Robinson, *loc. cit.*

can be calculated and applied for each measurement. The magnitude of the correction is seen from Table I where details are given of the values obtained under the worst conditions (*i.e.*, for the solution of highest conductivity) in each of the four systems. Except for the system phenol—water, the correction is very small and with this system, also, the value falls off rapidly with increasing concentration of phenol; the corrections for 50%, 60%, 70%, and 80% phenol are respectively 0.91%, 0.38%, 0.10%, and 0.03% of the observed values.

Table I—Correction for Conductivity

	Concentration of solution of maximum conductivity	Conductivity $\kappa$ ,	$C_1 + C_2$ ,	$C'_{obs} - C''_{obs}$ ,	Correction	Correction
System	phenol %	mho $\times 10^6$	$\mu\mu f$	$\mu\mu f$	$\mu\mu f$	%
Phenol—water	40.00	5.50	1500	334.2	6.18	1.85
Phenol— <i>m</i> -cresol	35.00	0.277	320	105.9	0.07	0.07
Phenol—aniline	82.50	0.417	320	106.1	0.17	0.16
Phenol— <i>p</i> -toluidine	80.01	0.670	320	104.1	0.43	0.41

In addition to its effect in introducing error directly into the capacitance observations, high conductivity of the dielectric, in producing a high equivalent added series resistance  $r$ , causes blurring of the resonance point and difficulty in its location; further, tighter coupling between the oscillator and the measuring circuit is demanded with consequent danger of feed-back effects. For these reasons it was found necessary, even with the large value of  $(C_1 + C_2)$  employed, to limit the measurements in the phenol—water system to solutions containing at least 40% of phenol.

**Results**—The concentration of each solution expressed as percentage by weight of phenol and also as molecular percentage of phenol, together with the corresponding value of the dielectric constant  $\epsilon$ , corrected for the conductivity of the solution, are given in Tables II, III, IV, and V for the systems phenol—water, phenol—*m*-cresol, phenol—aniline, and phenol—*p*-toluidine respectively. The dielectric constant is plotted against the percentage concentration by weight of phenol in fig. 4 and against the molecular percentage concentration in fig. 5.

### Discussion

Inspection of the curves in fig. 4 shows that they can be represented more accurately by a series of intersecting straight lines than by continuous curves. This does not imply that there is an abrupt discontinuity at the points of

Table II—Phenol—water, 70° C

Conc.	Mol. conc.	$\epsilon$
%	%	
40.00	11.32	27.67
44.91	13.50	26.72
50.01	16.08	25.13
55.00	18.97	23.67
59.99	22.31	22.64
65.00	26.23	21.37
69.99	30.87	19.94
75.00	36.49	18.23
80.01	43.39	16.36
85.00	52.04	14.41
89.62	62.31	12.81
94.98	78.37	10.94
100	100	9.03

Table III—Phenol—*m*-cresol, 50° C

Conc.	Mol. conc.	$\epsilon$
%	%	
0	0	9.32
9.98	11.30	9.41
19.99	22.30	9.51
25.00	27.69	9.53
30.00	32.99	9.59
35.00	38.22	9.60
40.00	43.37	9.64
44.98	48.43	9.72
50.00	53.47	9.76
55.00	58.41	9.80
59.98	63.26	9.86
65.00	68.09	9.87
70.02	72.85	9.95
75.01	77.52	9.97
79.97	82.10	10.00
80.00	82.13	10.03
84.99	86.68	10.07
90.01	91.19	10.17
100	100	10.28

Table IV—Phenol—aniline, 50° C

Conc.	Mol. conc.	$\epsilon$
%	%	
0	0	6.30
10.00	9.90	6.78
20.01	19.84	7.28
30.00	29.78	7.73
40.00	39.74	8.22
45.01	44.75	8.43
50.02	49.75	8.66
55.00	54.74	8.79
57.50	57.24	8.88

Table IV—(continued).

Conc.	Mol. conc.	$\epsilon$
%	%	
59.95	59.69	8.95
62.47	62.22	9.06
64.97	64.73	9.13
67.00	66.76	9.17
70.00	69.77	9.28
72.50	72.29	9.33
75.01	74.81	9.46
77.51	77.32	9.50
79.98	79.81	9.54
82.50	82.35	9.61
85.01	84.87	9.74
90.00	89.90	9.87
95.00	94.95	10.08
100	100	10.28

Table V—Phenol—*p*-toluidine, 50° C

Conc.	Mol. conc.	$\epsilon$
%	%	
0	0	5.07
10.00	11.23	5.68
20.00	22.15	6.29
30.00	32.79	6.91
40.00	43.15	7.57
47.50	50.74	8.04
50.00	53.23	8.14
52.50	55.72	8.31
55.00	58.18	8.47
57.52	60.65	8.63
60.00	63.06	8.73
62.50	65.48	8.89
65.00	67.89	8.96
67.50	70.27	9.01
70.00	72.65	9.07
72.50	75.01	9.13
75.00	77.35	9.20
80.01	82.00	9.43
90.00	91.11	9.83
100	100	10.28

intersection but only that the curves, by which the "linear" sections are united, are too slight to be detected.

This phenomenon is well known in liquids, such as concentrated sulphuric acid, in which chemical compounds are readily formed; but it becomes conspicuous only when co-ordinates are selected which give rise to approximately linear relations between composition and properties. Thus, in the present instance, if molecular percentages are employed, as in fig. 5, instead of weight percentages, the discontinuities are for the most part disguised. Indeed, it

is easy to show by drawing two intersecting ideal straight lines with weight percentages as abscissæ, as in fig. 4, and replotting with molecular percentages as abscissæ, as in fig. 5, that detection of the discontinuity then depends on the correctness of a single point at, or close to, the point of intersection instead of on the series of readings comprising the linear sections.

In most of the previous investigations on the dielectric constants of binary mixtures, determinations have been made at such wide intervals of compositions,

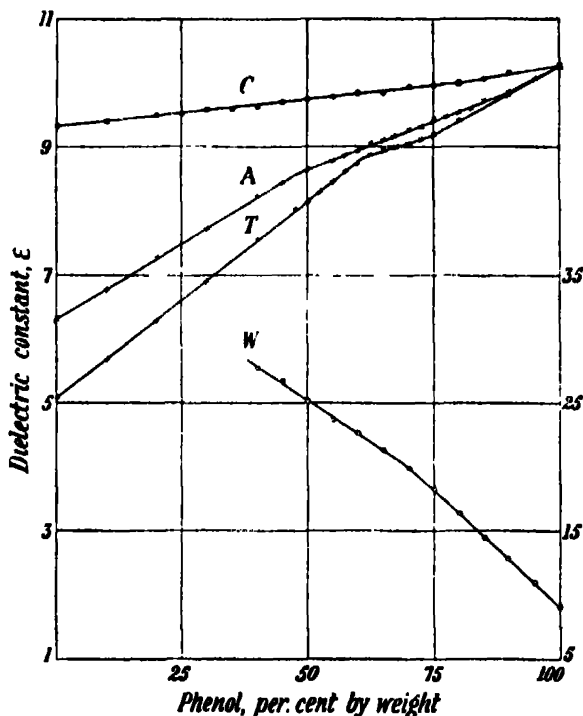


FIG. 4

W = phenol—water, scale on right  
 C = phenol—*m*-cresol  
 A = phenol—*a*niline  
 T = phenol—*p*-toluidine } scale on left

usually 20%, that the detection of discontinuities, such as are disclosed by the present measurements, would be impossible. Thereby, also, comparisons with the present measurements are unfortunately restricted.

The form of the curves cannot be explained on the assumption that the compounds indicated by the corresponding freezing-point curves are present in solution, since the discontinuities do not correspond to the compositions of the compounds or of the eutectics.



Thus curve W for phenol—water is best represented by two straight lines over the range examined, though the section on the water-rich side is perhaps not extensive enough to place its linear character beyond dispute. There is, however, obvious discontinuity at about 70% phenol, which does not correspond to the composition of phenol hemihydrate  $2\text{C}_6\text{H}_5\text{OH} \cdot \text{H}_2\text{O}$  (83.93% phenol) the compound clearly seen on the freezing-point curve of the system\* and crystallizing from solution.†

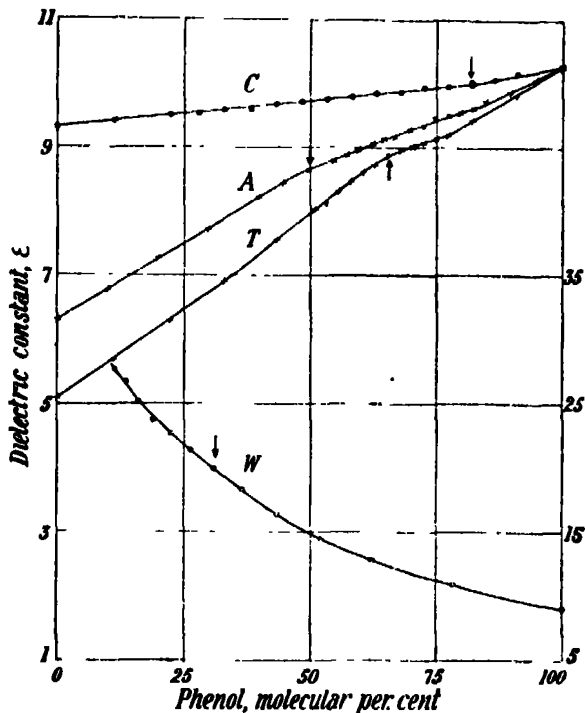


FIG. 5—References and scales as in fig. 4

Curve C for phenol—*m*-cresol consists of a single straight line over almost the whole of its course; it then rises more rapidly to 100% phenol, but the range of composition and the increase in dielectric constant are too small to establish its exact form. There is, however, a definite though small change of direction at about 80% phenol. The corresponding freezing-point curve‡

\* Rhodes and Markley, 'J. Phys. Chem.,' vol. 25, p. 527 (1921).

† Calvert, 'J. Chem. Soc.,' vol. 18, p. 66 (1865); Smits and Maarse, 'Verh. Akad. Wet. Amst.,' vol. 14, p. 192 (1911).

‡ Dawson and Mountford, 'J. Chem. Soc.,' vol. 113, p. 923 (1918); Fox and Barker, 'J. Soc. Chem. Ind.,' vol. 37, p. 268 (1918).

shows the formation of the compound  $C_6H_5OH \cdot 2CH_3C_6H_4OH$ , 30·32%

Each of the curves A and T for phenol—aniline and phenol—*p*-toluidine, respectively, consists of three distinct portions, which are essentially linear, intersecting at 45% and 90%, and at 60% and 75% respectively. The freezing-point curve of the system phenol—aniline\* shows the formation of the compound  $C_6H_5OH \cdot C_6H_5NH_2$ , 50·27% phenol, and that for phenol—*p*-toluidine† the compound  $C_6H_5OH \cdot CH_3C_6H_4NH_2$ , 46·76% phenol.

It has already been mentioned that the electrical conductivity—composition curves‡ for all these systems also give no indication of the existence in the liquid of the compounds which separate on freezing, but, on the contrary, indicate association in other molecular proportions. This has been attributed to inter-ionization of the constituents and the present results are not incompatible with this view. In particular, the intermediate portions of the curves for phenol—aniline and phenol—*p*-toluidine may represent the range over which the transition from the mode of ionization prevailing at the one end of the series to that at the other is predominating.

It may be observed that the base-rich sections of the curves for phenol—aniline and phenol—*p*-toluidine strike the ordinate for 100% phenol at very similar values, viz.,  $\epsilon = 11\cdot08$  and  $11\cdot27$  respectively; and similarly the phenol-rich sections strike the ordinate for 100% base at practically the same point, viz.,  $\epsilon = 5\cdot96$ , although the values for the bases themselves are considerably different, viz.,  $\epsilon = 6\cdot30$  and  $5\cdot07$  for aniline and *p*-toluidine respectively. Probably, therefore, it is the dielectric constant of the ion-pair produced on mixing which largely determines the mode of variation of dielectric constant with composition.

It is of interest to note that the variation in dielectric constant is evidently a very small factor in determining the change of electrical conductivity with composition for these binary systems, since in three of the four systems, over the range where the conductivity rapidly increases to a maximum, the dielectric constant decreases.

We desire to acknowledge the assistance of Mr. H. G. B. Robinson, M.Sc., and Mr. D. H. Thomas, B.Sc., in preparing the substances and making the measurements, respectively.

\* Schreinemakers, 'Z. phys. Chem.,' vol. 29, p. 577 (1899); Voano, 'J. Russa. Phys. Chem. Soc.,' vol. 48, p. 76 (1916).

† Kremann. 'Monatsh. Chem.,' vol. 27, p. 91 (1906).

‡ Howell and Handford; Howell and Robinson, *loc. cit.*

*Summary*

- (1) The dielectric constants of the binary systems phenol—water, phenol—*m*-cresol, phenol—aniline, and phenol—*p*-toluidine have been measured, the first at 70° C and the other three at 50° C.
- (2) The form of the curves showing the change of dielectric constant with composition is discussed.
- (3) There is no parallelism between the change of conductivity and that of dielectric constant with the composition of the mixtures.

---

*The Catalytic Properties and Structure of Metal Films. Part II.\*—  
The Electrical Condition of Platinum Films.*

By G. I. FINCH and (Miss) A. W. IKIN, Imperial College of Science and Technology.

(Communicated by W. A. Bone, F.R.S.—Received January 24, 1934.)

[PLATE 11.]

*Introduction.*

It was shown in Part I\* that platinum films formed by cathodic sputtering under suitable conditions were catalytically active in promoting the surface combustion of electrolytic gas at room temperature. Electron diffraction examination of such surfaces suggested that their catalytic powers were determined either by the nature of their structure and/or by the presence of oxygen occluded in the film during its preparation.

A series of investigations on the electrical condition of hot surfaces† had led to the view that the charge acquired by insulated metal surfaces was due, in the main, to the activation of adsorbed molecules of the gas or gases with which they were in contact; and, further, that such activation whereby the surface became electrically charged was in all probability a necessary condition for the occurrence of catalytic action.

\* Finch, Murison, Stuart, and Thomson, 'Proc. Roy. Soc.,' A, vol. 141, p. 414 (1933).

† Finch and Stimson, 'Proc. Roy. Soc.,' A, vol. 132, p. 192 (1931) and three earlier parts referred to therein, and Stimson, 'Proc. Roy. Soc.,' A, vol. 144, p. 307, Part V; Finch and Bradford, *ibid.*, p. 320, Part VI (1934).

The present experiments were carried out with the object of determining whether the property exhibited by certain thin films of platinum in promoting the union of hydrogen and oxygen at ordinary temperatures is determined by their structure or by the presence of oxygen occluded therein, or by both; and to what extent, if any, the electrical condition of such surfaces is associated with catalysis.

We have found, *inter alia*, that a relationship exists between the catalytic activity of sputtered platinum films and their electrical condition such that the rate of combustion of electrolytic gas in contact with the surface at room temperature is directly proportional to the rate of electrical charging up of the film; and, further, that the catalytic activity of such films is not determined by the presence of oxygen occluded therein during their preparation.

### *Experimental.*

*Preparation of the Platinum Films.*—Platinum films were prepared by cathodic sputtering on to 2 cm. diameter quartz or glass plates. In some experiments a sputtering vessel similar to that previously described was employed.\* Films free from oxygen, however, were deposited in the apparatus shown in fig. 1. The lapped brass surface of the anode, *a*, carrying the specimen plate, *p*, was sealed with low vapour pressure grease on to the glass sputtering vessel, *S*, which was fitted with the water-cooled platinum-faced cathode, *C*. The vessel, *S*, formed part of a closed gas-circulation system in which circulation was effected by means of the 3-stage mercury-vapour pump, *P*, discharging into the 3-litre reservoir, *V*, against a maximum pressure of less than 17 mm. The throttle system, *T*, served to control the rate of circulation; it consisted, in the main, of two throttles made from capillary tubing drawn out to suitable extents in a flame. *B*<sub>1</sub> and *B*<sub>2</sub> contained Ba-Mg alloy pellets welded to nickel supports. The auxiliary sputtering vessel, *D*, was fitted with a platinum wire cathode. All points, excepting those between the anode and the sputtering vessel, were fused glass or glass to silica sealed with low vapour pressure grease. The storage and sputtering vessels contained phosphoric oxide as shown.

After evacuation by means of a "Hyvac" pump backing a mercury vapour pump, the reservoir was filled to a pressure of about 10 mm. from the argon storage vessel, *A*, the throttle system and the tap, *t*, being closed and the mercury-vapour pump in full operation. The argon was supplied by the

\* Finch, Murison, Stuart and Thomson, 'Proc. Roy. Soc.,' A, vol. 141, p. 414 (1933).



FIG. 3.—(100), also (112) parallel to film. FIG. 4.—(111), also (100) and (112) parallel to film.

FIGS. 3 and 4.—Electron diffraction patterns from catalytically inactive and electrically neutral platinum films sputtered in argon on to quartz discs.



FIG. 8.—Very active.



FIG. 9.—Moderately active.



FIG. 10.—Almost inert.

FIGS. 8 to 10.—Electron diffraction patterns from catalytically and electrically active films.



FIG. 11.



FIG. 12.



FIG. 13.

FIGS. 11 to 13.—Patterns showing effect of heating on crystal structure of Pt. films.

FIG. 11, after 8 mins. heating; FIG. 12, after 16 mins. heating; and FIG. 13 after 2 hours heating.



British Oxygen Company and was stated by them to be 100% pure. The Ba-Mg alloy pellets in  $B_1$  and  $B_2$  were flashed by high frequency eddy-current heating on to the walls of the two vessels which were evacuated by the pump, P, the throttle system being closed and the tap,  $t$ , open. The throttle was then adjusted to permit of a suitable rate of circulation of argon throughout the apparatus, the pressure in S being observed on the McLeod gauge. During circulation, platinum was sputtered vigorously in the vessels, D and S. This purification process, which served to eliminate all impurities other than rare

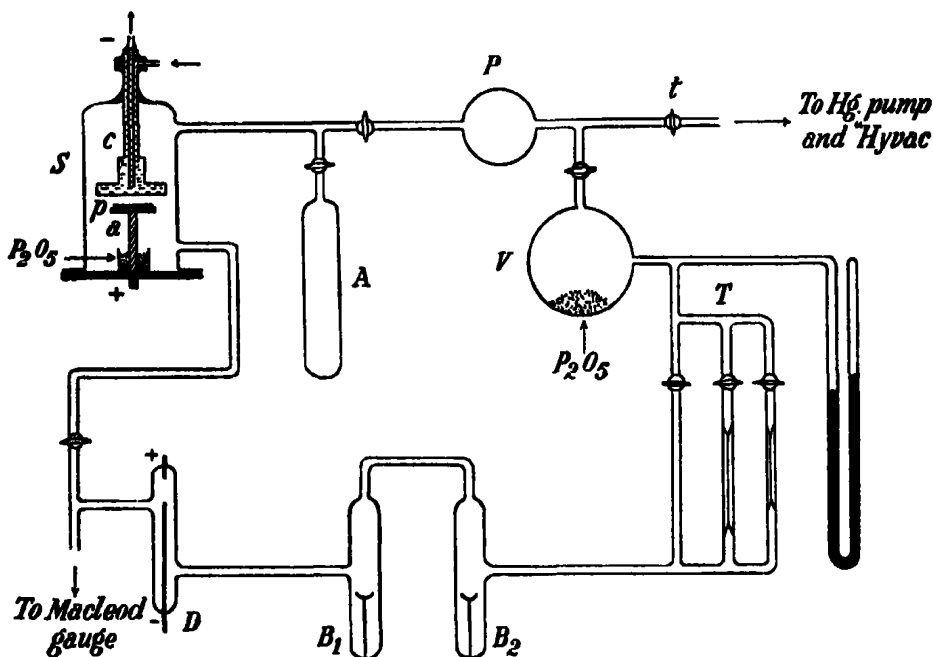


FIG. 1.

gases, was carried out until several complete circulations had been effected. The necessary stop-cocks, as shown in fig. 1, were then closed in order to permit of the introduction of a specimen plate into the sputtering vessel, and its subsequent evacuation, without introducing impurities into the purified argon. Sputtering was then carried out in the normal manner, argon being, however, continually circulated at the desired pressure.

*Measurement of the Rates of Charging Up of the Platinum Films.*—The electrical measurements and the determination of catalytic activity were carried out in the apparatus shown in fig. 2, into which the film was transferred

immediately after preparation. The vessel, *V*, was lined internally as shown by a broken line, with a thick, practically opaque film of sputtered silver connected to earth through the fused-in platinum leads,  $l_1$  and  $l_2$ . The platinum wire, *w*, supporting the film, *S*, under examination, was led to a Lindemann quadrant electrometer, the necessary insulation being provided by two quartz tubes, *q*, with vacuum-tight, sealing-waxed joints suitably disposed in order to ensure a long surface-leakage path, as shown at *g*. The electrometer needle could be charged to  $\pm 1.019$  volts by means of a silver chain swung into contact with the wire, *w*. The quadrant pairs were across a suitable

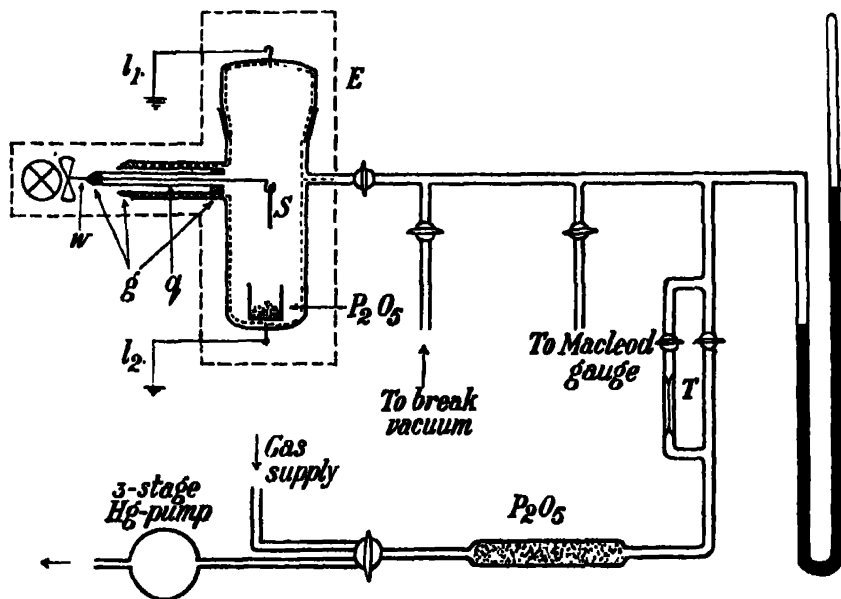


FIG. 2.

accumulator, connected to earth through a  $0.5 \text{ M}\Omega$  bridging potentiometer. An earthed sheet-metal case, *E*, enclosed the electrical system. The arrangement outlined permitted of rapid and accurate adjustment into coincidence of the electrical and mechanical zeros of the electrometer; and also enabled its sensitivity, needle-film system insulation, and freedom from spurious charges to be determined quickly. The capacity of the needle-film system to earth was constant throughout all experiments at 80 cm. Fig. 2 is self-explanatory of the services enabling the catalytic activity of the film suspended at *S* to be determined. The throttle, *T*, was found convenient in enabling close adjustment of gas pressure to be obtained. Electrolytic gas, prepared



by electrolysis of recrystallized barium hydroxide, was dried over purified phosphoric oxide before admission to the apparatus.

*Electron Diffraction Examination.*—The examination of the structure and of certain other properties of the films was carried out in the electron-diffraction camera previously described.\* Special care was taken throughout in handling to avoid soiling the films, particularly with traces of grease. Such precautions were necessary because it was found that films sputtered in argon were extremely sensitive to poisoning, in the catalytic sense, particularly by grease vapours.

### *The Results.*

The catalytic activity, measured in terms of the rate of combustion of electrolytic gas at room temperature, and the electrical condition of 65 platinum films deposited on 2 cm. diameter glass or quartz discs by cathodic sputtering from two different samples of platinum were determined. In addition, 160 electron diffraction patterns were photographically recorded and analysed. For present purposes it will suffice, however, to give a general summary of the results obtained together with a somewhat fuller account of a few typical examples. It will be unnecessary to set forth in detail the sputtering conditions employed in the production of the different types of films examined, because these have been given elsewhere.† The following results were obtained :—

*Catalytically Inactive Films.*—Without exception, those sputtered platinum films which proved to be catalytically inactive also remained uncharged and failed to acquire a surface potential either *in vacuo* or when in contact at room temperature with either electrolytic gas, hydrogen, or oxygen. In all, 25 such electrically neutral and catalytically inert films were examined. All gave rise to electron-diffraction patterns which showed that their structure was that of normal face-centred cubic platinum,  $a = 3.91$  Å. In addition, a greater or lesser proportion of the crystals in each film were orientated in certain preferred directions with respect to the film surface. The patterns, figs. 3 and 4, Plate 11, and patterns similar to those previously published by G. P. Thomson and his co-workers‡ and by Finch and Quarrell,§ showing pronounced arcing of the diffraction rings are representative of those obtained from the different

\* 'Proc. Roy. Soc.,' A, vol. 141, p. 398 (1933).

† 'Proc. Roy. Soc.,' A, vol. 141, p. 414 (1933).

‡ 'Proc. Phys. Soc.,' vol. 45, p. 381 (1933).

§ 'Proc. Roy. Soc.,' A, vol. 141, p. 398 (1933).

types of catalytically and electrically inert films, and enable a better estimate of the relative proportions of orientated and randomly disposed crystals to be made than could otherwise be drawn from a purely descriptive account. It will be seen that as many as three distinct types of orientation were observed for a single film, fig. 4. In no case of inactivity was a pattern of complete rings alone, *i.e.* one indicative of a wholly random crystal array, obtained. A wide range of film thicknesses was examined, varying from semi-transparency to complete opacity.

*Catalytically Active Films.*—Forty platinum films exhibiting a wide range of catalytic activity, equivalent to between  $0.28 \times 10^{-2}$  and  $285.0 \times 10^{-2}$  c.c.

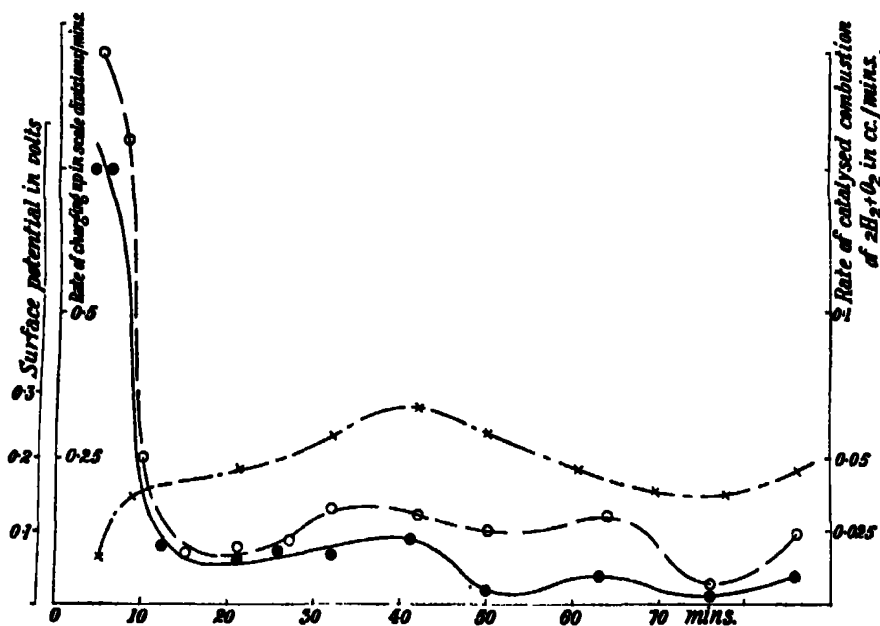


Fig. 5.—Capacity of surface-needle system = 80 cm. 1.019 volts = 31.3 scale divisions.  
 × surface potential; O rate of charging up; ● rate of combustion.

of electrolytic gas at N.T.P. burnt to steam per minute, were obtained and examined. The films always acquired a surface potential when in contact with, and promoting the combustion of, electrolytic gas; and, furthermore, the rate of build-up of such potential was closely paralleled by the rate at which the film promoted the combustion of electrolytic gas. In 22 experiments the charge was positive throughout the active life of the film; in 10 others, during the initial stages of activity the surface acquired a negative

potential which, however, at a later stage, coinciding usually with the complete development of full activity, changed sign to positive. In the cases of the remaining eight films the charge proved to be negative throughout the entire active life of the film as a catalyst. These facts may be illustrated by the results shown in figs. 5, 6, and 7, in which the changes in surface potential, catalytic activity, and the rate of charging up during the active lives of characteristic examples of these three different types of films are graphically recorded.

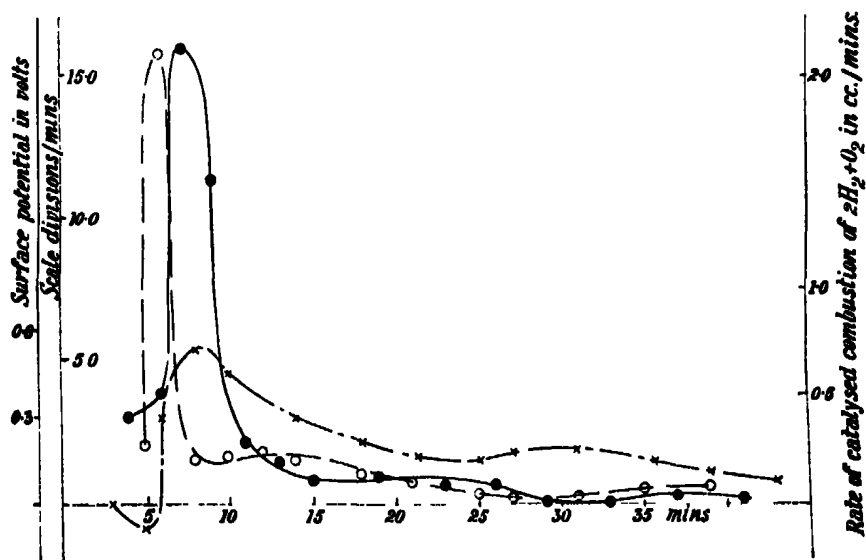


FIG. 6.—The apparent time lag between the charging up and combustion rate curve is in effect due to the delay involved in taking the gas pressure readings.  $\times$  surface potential;  $\circ$  rate of charging up;  $\bullet$  rate of combustion.

Twenty-eight active films were examined by electron diffraction. All exhibited a crystalline structure characteristic of normal platinum; in twenty films the crystal array was completely random, and they also proved to be both catalytically and electrically more active than eight other films, prepared with high sputtering currents and a relatively hot cathode, which showed more or less pronounced signs of crystal orientation, with either (100), (110), or (111) faces parallel to the film surface. In general, it was found that the stronger the orientation, the less active was the film as a catalyst. Thus the film yielding the pattern, fig. 8, Plate 11, and in which the array of relatively small crystals (as deduced from the width of the diffraction rings) was wholly random,

exhibited great activity, equivalent to as much as  $217.0 \times 10^{-3}$  c.c. of electrolytic gas at N.T.P. burnt to steam per minute. Another film, fig. 9, Plate 11, the pattern from which showed faint signs of orientation, was moderately active ( $11 \times 10^{-3}$  c.c.), whereas the film, fig. 10, Plate 11, containing a fairly large proportion of orientated crystals (111) parallel to film) was only barely perceptibly active.

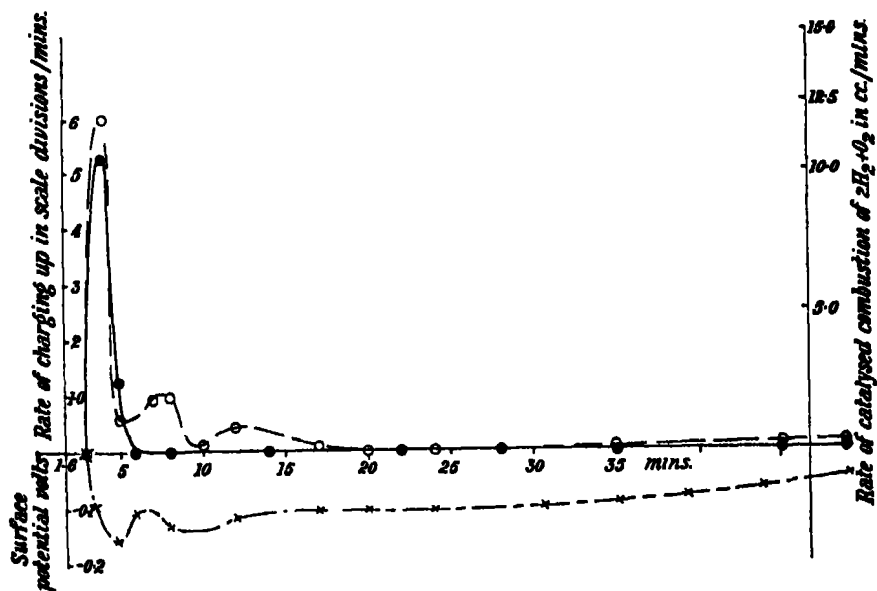


FIG. 7.—× surface potential; O rate of charging up; ● rate of combustion.

For convenience in comparison, the direction of the current has been neglected in deriving the rates of charging.

**Effect of Heat on the Films.**—In Part I it was shown that heating an argon sputtered platinum film to  $240^{\circ}\text{C}$ . destroyed its potential activity as a catalyst. We have now confirmed this observation in further experiments and in so doing have also extended the temperature conditions with surprising results. It was found that, whilst heating for a time which amply sufficed to destroy all traces of latent activity had no appreciable effect upon the appearance of the diffraction pattern, a suitably prolonged and intensive heating led eventually to an apparently complete breakdown of crystal structure. Thus in one such experiment a film, the counterpart of which was exceedingly active on development in electrolytic gas, was heated *in vacuo* by radiation from a 1 cm. distant dull-red platinum filament during  $\frac{1}{2}$  minute, whereupon the film, after cooling, was found to be completely inactive. Two minutes after resumption of heating

a diffraction pattern was recorded which did not differ appreciably either as to background or ring sharpness or, indeed, in any other material respect from the pattern, closely similar to fig. 8, Plate 11, obtained prior to heating. In this and in other experiments of a similar nature heating under the conditions outlined above was continued. After 8 minutes the pattern, fig. 11, Plate 11, was photographed, after a further 8 minutes, fig. 12, Plate 11, and, finally after 2 hours heating in all, fig. 13, Plate 11. The broadening and progressive loss of definition of the rings leading to their eventual disappearance is evidence to the effect that the prolonged heating had resulted in the gradual breakdown of the crystal structure into the amorphous state.

### *Discussion.*

Several attempts have been made in the past to determine a relationship between thermionic emission from a surface and its catalytic activity. Thus Brewer\* found that emission from the catalysing surface of both positive and negative charges occurred in a number of heterogeneous gaseous reactions, but failed to obtain a quantitative relationship between such emission and reaction velocity. Similarly, Bradford,† in this laboratory, studied the emission from a hot platinum wire promoting the combustion of carbonic oxide detonating gas, but was unable to establish any wide proportionality between the rate of such reaction and the thermionic emission. On the other hand, the results of a series of studies on the electrical condition of hot surfaces,‡ pointed towards the existence of a close connection between their electrical and catalytic properties. For example, it was found that with increasing temperature the specific effect of the surface upon its equilibrium potential diminished until, at a sufficiently high temperature, the value of the potential was characteristic of the gas alone and practically independent of the nature of the surface. Furthermore, in recent communications to the Society by Bradford and one of us, and by Stimson, it has been shown that the rates of charging up of a number of different surfaces in contact with various gases increased with temperature in an exponential manner such that a rapid transformation from relatively low to high rates of charging up took place in that temperature region within which the surfaces commenced to exhibit a similar change in

\* 'J. Phys. Chem.,' vol. 32, p. 1008 (1928).

† 'J. Chem. Soc.,' p. 1544 (1932).

‡ 'Proc. Roy. Soc.,' A, vol. 144, p. 320, Part VI (1934), and earlier parts referred to therein.

catalytic activity. In connection with this it may be recalled that Bone,\* in 1914, first drew attention to the fact that the specific effect of various surfaces in promoting combustion tended to disappear with increasing temperature, no matter how widely they might otherwise have differed in this respect at lower temperatures, an observation which thus affords a striking parallel with the changes in the electrical condition of hot surfaces with temperature, such as we have previously found to occur. Finally, the experiments described above have shown clearly and beyond all reasonable doubt that fluctuations in the rates of electrical charging-up of a cool platinum film promoting the combustion of electrolytic gas are accompanied by corresponding changes in the reaction velocity and hence in the catalytic activity of the surface.

It has previously† been shown that a suitably activated platinum sheet acquired a positive charge, even at ordinary temperatures, when in contact with either oxygen or hydrogen alone. In view of this fact the remarkable relationship, now shown to exist between the electrical and catalytic activities of sputtered platinum films while promoting the union of electrolytic gas at room temperature, enables the conclusion to be drawn that the act of catalysis is both preceded and determined by an interaction between the surface and one or both constituents of the combining mixture, whereby the catalyst becomes electrically charged.

The results of experiments carried out with different metallic surfaces by Stimson and one of us,‡ and more recently by Bradford and one of us§ with gold, led to the view that both the catalytic and electrical properties, which have now been shown to be so closely interrelated, were themselves in some manner intimately associated with, and determined by, the nature of the structure of the catalyst; and, the results outlined in Part I of this series supported this view, though they did not enable the precise nature of this association to be determined. The results of the structural examination set forth above together with those of Part I, though not affording a complete answer to this question, do, however, enable the issues to be narrowed down.

The diffraction patterns obtained can be divided into three main classes according to whether the rings (i) are of uniform circumferential intensity; (ii) are broken up into relatively short arcs; or (iii) consist of short, more or less intense arcs superimposed upon rings of lesser but otherwise uniform

\* 'Roy. Soc. Arta.,' Howard Lectures (1914).

† Finch and Stimson, 'Proc. Roy. Soc.,' A, vol. 124, p. 356 (1929), Part III.

‡ 'Proc. Roy. Soc.,' A, vol. 116, p. 379 (1927).

§ 'Proc. Roy. Soc.,' A, vol. 144, p. 320 (1934).

circumferential intensity. In structures producing rings of type (i) the crystals are randomly disposed; arcs as in (ii) show that the crystals are orientated, either all in one preferred direction, or some in one and others in another. Arcs and uniform rings together testify to the orientation of some crystals and the disarray of the remainder.

Films containing only orientated crystals were, without exception, completely inactive, whilst in films of pronounced activity the crystals were randomly arrayed. Nevertheless, these facts do not justify the supposition that orientation is, *per se*, inimical to activity, because films containing both randomly disposed and orientated crystals were also as a rule completely inactive, though in some cases where the proportion of orientated to disarrayed crystals was low, slight activity was observed. Moreover, as has been shown above experimentally, the activity of a film with disordered crystal array could be completely destroyed by gentle heating without in any way affecting the appearance of the diffraction pattern. Thus it would seem that the inertness of films containing orientated crystals is due to sputtering conditions leading to orientation being either destructive of active films, or inimical to their formation. These facts also negative any suggestion that crystal size might be directly associated with activity, because heating did not affect the width of the rings until long after complete destruction of latent activity. It would, therefore, appear that the structural features determining the activity of sputtered platinum films in promoting the combustion of electrolytic gas at room temperature are of such a nature as not to be directly revealed by electron diffraction analysis.

In Part I it was found that films rapidly sputtered under cool conditions were far more active catalysts than slowly deposited films of approximately similar thickness, and this observation has been fully borne out in the present experiments. Furthermore, inspection of the 112 patterns which we have now obtained from potentially active platinum films has shown that the more active the film, the greater is the amount of background due to general electronic scattering appearing in the corresponding diffraction pattern. The glass or fused quartz receivers employed in these experiments did not yield any diffraction pattern, but simply gave rise to random scattering, except so far as charging-up effects resulted in unstable localizations of electrons. The sputtered films deposited on these receivers were, however, in general far too thick to permit of the substrate contributing to any random scattering in addition to that produced by the film itself. The general diffuse background visible in our diffraction patterns can therefore be wholly attributed to the platinum films.

Inelastic scattering, scattering by amorphous arrays of atoms and multiple elastic scattering can all give rise to background, and these factors have not been under control in our experiments. Nevertheless, the fact that background intensity increased with catalytic activity seems to be of special significance, because in Part I it was established that (i) high sputtering rates under cool conditions yield much more active films than do low rates under otherwise similar conditions, and (ii) crystal size diminishes with increasing rate of sputtering. Furthermore, since it is now well known that metals are sputtered as discrete atoms, it seems reasonable to conclude that rapidly sputtered films will contain more atoms not entering into any regular structure than a slowly sputtered film.

The facts and considerations outlined above suggest that the potentially active catalyst centres may possibly consist of platinum atoms not forming part of any ordered array or crystal structure. It is true that activity could be destroyed by heating without affecting the background intensity; but changes in the density of packing together of such isolated atoms, such as might well be brought about by heat, are hardly likely to affect the background, provided such density changes do not at the same time result in the atoms entering into ordered array. In connection with this it is significant that, as has now been shown experimentally, suitable heating of a potentially active polycrystalline platinum film at first rapidly destroyed such activity, and then gradually broke down the crystalline structure to an amorphous and catalytically still inactive state; whilst development in electrolytic gas followed by destruction of activity by drying led to a pronounced clearing-up of background and increase in crystal size.\*

One of us (G. I. F.) wishes to thank the Government Grant Committee of the Royal Society, Imperial Chemical Industries, Ltd., and Messrs. Ferranti, Ltd., for grants which defrayed the cost of much of the equipment employed.

### *Summary.*

The surface potentials and rates of charging-up of cathodically sputtered platinum films in contact with electrolytic gas at room temperature have been determined, and the film structures examined by the method of electron diffraction.

It has been found *inter alia* that (i) electrically active films were also catalytically active, whilst electrically neutral films were inert as catalysts; (ii) the

\* *Loc. cit.*, Part I.



rates at which active films promoted the union of  $H_2$  and  $O_2$  were directly proportional to their rates of charging-up; (iii) with the more active films, the crystals were arrayed at random, whereas with inactive films a considerable proportion of crystals possessed a common direction of orientation; (iv) the diffuse background in electron diffraction patterns obtained from active films increased in intensity with increasing activity; (v) heating rapidly destroyed the latent activity of a film and when sufficiently prolonged also broke down the crystal structure.

From these facts, in conjunction with those set forth in Part I, it is concluded that, under the conditions of these experiments, (i) catalytic action was determined by a prior interaction between the surface and one or both constituents of the combining mixture, whereby the catalyst became electrically charged; (ii) activity was not determined by either crystal size or orientation; (iii) catalytic activity may have been centred round atoms of platinum not forming part of any ordered array or structure.

---

### *A Quantitative Study of Pleochroic Haloes.—I.*

By G. H. HENDERSON, Ph.D., King's College, and S. BATESON, M.Sc., Dalhousie University, Halifax, N.S.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received January 29, 1934.)

[PLATES 12-14.]

#### *Introduction.*

The beautiful phenomenon of pleochroic haloes is recognized as one of the most striking manifestations of radioactivity. The nature of these haloes was first explained by Joly in 1907 and much research has confirmed and expanded his theory.\*

These haloes consist of minute discolorations occurring in certain minerals, notably mica (biotite). They are spherical in shape and bear at the centre a very small inclusion or nucleus. A thin section of one of these spheres shows under the microscope a series of concentric dark rings, whose radii may be

\* For a very complete summary see the article by Holmes in "The Age of the Earth," 'Bull. Nat. Res. Council No. 80,' Washington, 1931.

identified with the ranges of  $\alpha$ -particles ejected from radioactive substances in the nucleus. The individual haloes present varied appearances, but may be classified according to their blackening in a regular order of development.

The accepted theory of halo formation assumes that the nuclei in a crystal of mica were all formed simultaneously. These nuclear inclusions contained uranium which disintegrated with the emission of  $\alpha$ -particles. In the course of a relatively short time the whole uranium family came into equilibrium and then continued to emit  $\alpha$ -particles throughout many million years. These particles, passing through the surrounding mica, gradually altered it chemically as evidenced by discoloration. Corresponding to the different radioactive content of the nuclei the surrounding regions were exposed to different amounts of radiation and hence there are found to-day in the same cleavage flake haloes in all stages of development.

A peculiar interest attaches to these haloes because the darkening must have proceeded slowly throughout geological time. Hence they may furnish evidence regarding the origin of the different radioactive series and of the minerals in which they occur. Moreover, the formation of a halo is a very delicate test of the presence of a radioactive element, surpassing in sensitiveness any other method of detection. The study of haloes thus offers the possibility of deducing the existence of hitherto unknown elements, which perhaps have long since vanished from the earth.

It is advisable that further study of pleochroic haloes should be carried out and that new methods of attack on the problem should be devised. The method usually adopted is visual and subjective. The radii of the rings are measured with a micrometer microscope and the sequence of halo development followed by a visual estimation of their density and structure. While much information has been gained by this method, subjective measurements are liable to be in error because of the readiness with which the eye may be deceived. Moreover, this method does not allow of more than a very qualitative explanation of the processes involved in the formation of a halo. The first step towards a more quantitative method was taken by Kerr-Lawson,\* who took photomicrographs of the haloes and then measured the density variations with a Moll recording photometer. The records showed clearly the main features of the haloes, but the method is open to objections. First, the record of blackening in the halo may be distorted unless the calibration curve of the photographic plate is considered, and second, there is the additional distortion due to any flaws which may occur in the plate.

\* "Univ. of Toronto, Geol. Studies Nos. 24 and 27" (1927).

It was felt that a more direct measurement of the blackening in a halo was desirable. In order to do this it was necessary to design a recording microphotometer capable of measuring accurately the variations in blackening across an object only 0.007 cm. in diameter. It is the purpose of this paper to give a short description of this apparatus and to discuss the results thus far obtained with it.

### *The Halo Photometer.*

Full details of the apparatus, which may be called a halo photometer, will be published elsewhere, consequently only a general description need be given here. A microscope A, fig. 1, is mounted very rigidly with its axis horizontal.

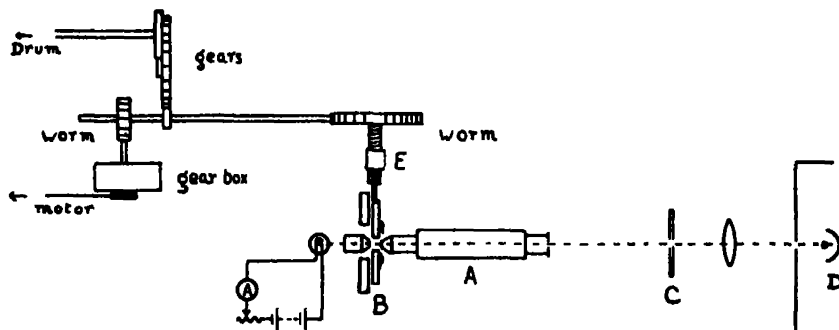


Fig. 1.

To its stage are fixed guides which permit a carriage B to move horizontally at right angles to the axis of the microscope. The cleavage flake of biotite containing the halo to be photometered is mounted on a slide and clamped to the carriage B. An image of the halo is projected on the slit C with the aid of an automobile headlight lamp and a condensing lens. The magnification is about 200, the diameter of the image being about 1.4 cm. The slit opening is about  $1.8 \times 0.05$  mm. which should afford ample resolving power. It permits light from this small section of the image to pass into a lens which focusses it on a vacuum caesium photoelectric cell D.

The current from the cell is amplified by means of a single tube (UX 232) screen grid resistance coupled amplifier having an amplification factor of  $8.5 \times 10^5$ . The output current of the amplifier is recorded by a d'Arsonval galvanometer. A change in photoelectric current of the order of  $10^{-14}$  ampere produces a deflection of 1 mm. of the galvanometer. The amplifier characteristics were so arranged that a linear relationship existed between input and output currents. The carriage B holding the halo is made to move slowly and

uniformly across the microscope stage by means of a nut moving along a carefully made micrometer screw E. The motion was slow enough so that the critically damped galvanometer might follow faithfully the changes in plate current due to the variations of light intensity falling on the cell as the halo travelled across the slit. The screw is also geared to a drum bearing bromide paper on which is recorded the spot of light from the galvanometer. Thus as the halo slowly passes across the slit, the drum turns regularly in conjunction with it and receives the record of the varying intensity of the light transmitted through the section of the halo picked out by the slit. The optical system of the microscope consisted of a 4 mm. Spencer apochromatic objective and an  $8\times$  Leitz Periplan ocular. The screw E was driven at the rate of one turn in 100 minutes by an electric motor through a system of gears with a reduction of about 180,000 to 1. About 10 minutes were required for the recording of a halo.

The photoelectric cell and amplifier were tested for proportionality of response to light intensity as follows. A small flash light lamp was placed at different distances from the slit so as to give about the same range of light intensities as that met with in the haloes. The deflections of the galvanometer were proportional to the inverse square of the distance within the limits of experimental error of about 1%.

It is, of course, absolutely necessary to ensure that distances on the halo are proportional to distances on the bromide paper and to know the enlargement factor. This was done by taking a record of a microscope test plate with rulings 0.01 mm. apart, mounted on the carriage in place of the mica. A photograph of such a record is shown in fig. 2, Plate 12. The light intensity diminishes towards the top of the photograph so that the longer traces at the bottom correspond to the transparent parts of the scale and the peaks above to the opaque rulings. It will be seen that the recording is very regular. The average deviation from the mean of the distances between successive peaks is less than 1% and even this may be due to irregularities in the scale itself rather than in the mechanism. On the original of fig. 2 the mean distance between the peaks is 19.7 mm. corresponding to 0.01 mm. on the test slide. It was arranged that the same portion of the screw should always be used so as to avoid any periodic error in the screw. During the operation of the halo photometer over an extended period the enlargement factor gradually increased by about 4%. Test slide records were taken frequently and it is probable that the error in the enlargement factor applicable to any particular record is less than 1%.

The use of the test slide also permitted a test of the resolving power of the photometer. This was done by comparing the record with the theoretical curve to be expected from a slit of finite width passing over an opaque line. When allowance was made for irregularities at the edge of the ruling, the two curves agreed very well.

#### *Nature of Specimens Used.*

Two specimens of biotite will be dealt with in this paper. One was a brown biotite from Mount Apatite, Maine, and the other a green biotite from near Murray Bay, Quebec. The latter, kindly furnished by Professor T. L. Walker, of Toronto, was a sample of the same material used by Kerr-Lawson and thus permitted accurate comparison with his results. Both micas contained a moderate number of uranium haloes. A number of other biotites, chiefly of Canadian origin, were examined but proved blank.

The thickness of cleavage flake which was used varied considerably. When the haloes showed fine detail a very thin section containing the nucleus was used so as to avoid loss of detail due to overlapping by non-central sections. In such cases the thickness was of the order of  $5\ \mu$  or about  $1/14$ th of the diameter of the halo sphere. When great detail was not apparent a thickness of twice or even three times this amount was used.

When the nucleus of the halo was clearly visible there was no difficulty in distinguishing a central section. In most cases, however, the nucleus was too small to be distinguished, but after some experience the choice of a central section was not difficult.

The various stages of development observed in the microscope have often been described and need not be repeated here. We have found it convenient to classify our haloes into four groups as follows :—

*Embryonic*.—Containing all haloes at an early stage, including faint indications of the outer or RaC' ring.

*Normal*.—In which the ring structure is most clearly defined.

*Overexposed*.—In which the blackening has proceeded so far as to obliterate most of the inner structure.

*Reversed*.—In which clear signs of " reversal " are shown.

Embryonic and over exposed haloes proved to be the most common.

#### *Description of Records.*

A series of records of Mount Apatite haloes in progressive stages of development is given in figs. 3-9, Plate 12, and a similar series for Murray Bay in figs.

10-17, Plate 13. The haloes from the two micas show the same general features, but also certain characteristic differences to be discussed later. A few photomicrographs are shown in figs. 18-25, Plate 14, which give an idea of the type of halo producing the various records. They were taken with the same optical system as was used in the photometer.

On the records, Plates 12 and 13, the abscissæ represent distances along a diameter of the halo. The scale varies slightly from record to record on account of the slowly changing enlargement factor. The ordinates are galvanometer deflections and are proportional to the light absorbed by a small area of the halo. The intensity of transmitted light diminishes towards the top of the record. The scale is arbitrary, depending on the thickness of the biotite flake and the amplification factor used.

The records in Plates 12 and 13 have been chosen from a number of photometer records in order to illustrate what we think to be the various stages of development. It has not been possible to obtain a complete series of haloes from the small specimens of biotite available. The continuity seems clear, however, except for the embryonic stage which is somewhat uncertain and will be referred to later.

A few explanatory remarks may be made about individual records. It is difficult to obtain the haloes free from flaws, such as cracks in the mica. Once the halo photometer is working properly, the main difficulty encountered in practice is the preparation of flakes sufficiently flawless to give good records. The effect of a system of cracks is seen just to the right of the centre of fig. 4, Plate 12. Although they spoil the symmetry of the record, these cracks need not necessarily interfere with the measurement of some features of the halo.

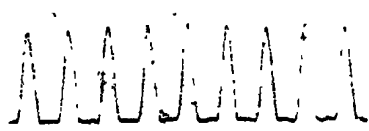
The record of a very perfect normal Murray Bay halo is shown in fig. 12, Plate 13. It shows clearly the five rings described by Kerr-Lawson, including the innermost  $r_0$  ring first discovered by him. A sixth ring may also be distinguished just inside the outer ring. This ring, which would be noted as  $r_4$  in Kerr-Lawson's notation, will later be shown to be due to the actinium series.

On fig. 6, Plate 12, are given two records of the same over exposed halo in different azimuths which happened to be taken on the same sheet. Fig. 9, Plate 12, is a record of a reversed halo from Mount Apatite. It is difficult to secure good thin sections of such haloes as the mica in the halo is so altered that the cleavage rarely continues smoothly through the halo. The resulting cracks produce the irregularities seen at both ends of the record. A very interesting reversed halo from Murray Bay is shown in fig. 17, Plate 13. The

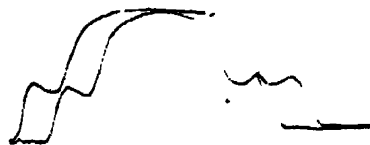








2



6



3



7



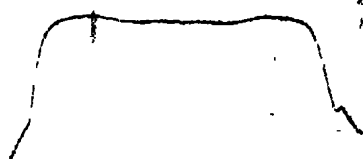
4



8



5



9



**10**



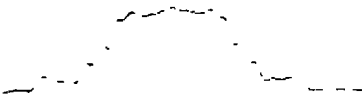
**14**



**11**



**15**



**12**



**16**



**13**



**17**



18



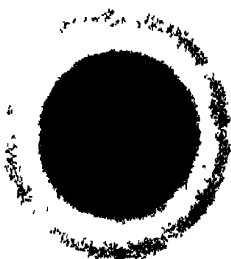
19



20



21



22



23



24



25



minima are quite characteristic and are almost certainly due to reversal. Both these reversed haloes have unusually large diameters, probably due to large nuclei.

### Measurement of Records.

It is advisable first to consider the distances from the centre of the various singularities of the records. This gives a comparison with the ranges of the various  $\alpha$ -particles involved as well as with the measurements of ring radii of previous investigators.

The singularities of a halo are most clearly shown in the normal type, many of them being missing in other stages. We have found it convenient to adopt the system of lettering given in fig. 26 to denote these features. This figure is a composite tracing of the records of normal haloes given in figs. 5 and 12.

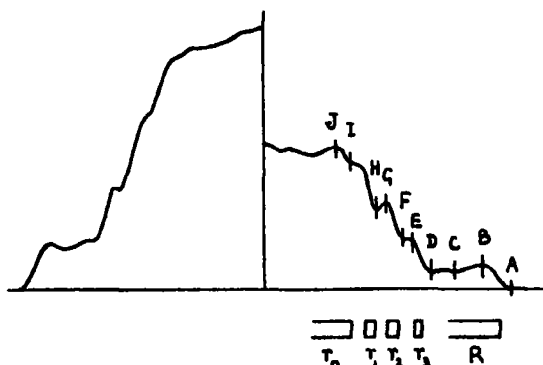


FIG. 26.

Underneath the tracing are shown the rings observed in the microscope by Kerr-Lawson. Thirty records were measured with the results given in Table I. Measurements were made with a transparent scale, being converted into microns with the aid of the test plate records. It will be seen that there is good agreement among the first 20 records. The average deviations from the means are given at the foot of the table. Some features such as A are measurable with greater accuracy than others as they are more definite. There seems to be no appreciable difference between the two micas. For instance, the mean radius of the outer ring A is  $34.45 \mu$  for Mount Apatite and  $34.38 \mu$  for Murray Bay.

Ten of the records seemed anomalous and were excluded from the means. Of these, five showed radii much lower than the means. This discrepancy is probably due to the flake being off centre. This was verified for Nos. 66 and

Table I.—Radii in microns.

No.	Description.	A.	B.	C.	D.	E.	F.	G.	H.	I.	J.
40	Mount Apatite— Normal	34.8	31.6	28.6	—	21.1	20.3	17.3	15.9	—	—
62	Over exposed	34.6	28.5	—	—	—	—	—	—	—	—
67	"	34.3	—	—	22.7	—	—	—	—	—	—
71	Embryonic	34.3	—	—	23.7	20.6	19.4	—	14.7	—	—
72	"	34.5	—	—	—	—	—	—	—	—	—
70	Over exposed	34.5	—	—	—	—	—	—	—	—	—
47	Embryonic	34.0	—	—	22.6	—	19.3	—	14.8	—	—
28	Normal	34.6	30.8	27.9	23.7	21.5	20.7	17.1	16.2	11.3	9.5
54	Murray Bay - Normal	34.8	30.8	28.2	23.4	20.5	19.6	16.9	15.8	12.6	9.3
60	"	34.8	31.2	26.9	23.6	20.4	19.7	16.9	15.8	12.6	9.6
65	Over exposed	34.4	29.4	26.0	22.7	—	—	—	—	—	—
76	Normal	34.4	30.8	—	23.5	20.7	19.4	16.3	15.2	—	—
79	"	34.0	—	—	—	—	—	—	—	—	—
86	"	33.8	30.1	26.6	22.8	20.2	19.0	16.6	15.2	12.7	9.5
93	"	34.8	31.3	27.6	23.7	21.1	20.0	17.5	16.2	12.8	—
95	"	34.5	—	26.8	23.1	20.6	19.6	17.2	15.9	—	—
94	"	34.8	—	26.2	23.4	20.4	19.8	17.2	16.0	—	—
91	"	34.0	—	—	22.8	19.8	19.0	16.4	15.2	—	—
92	Over exposed	34.3	—	—	—	—	—	—	—	—	—
96	Normal	34.0	—	—	22.4	—	—	—	—	—	—
66	Murray Bay— Normal	33.5	—	—	21.8	18.8	17.4	15.5	13.4	—	—
74	"	33.2	—	24.5	19.8	17.6	15.7	12.8	—	—	—
89	Over exposed	33.5	—	—	20.3	—	—	—	—	—	—
78	Normal	33.4	—	—	21.7	18.0	—	—	—	—	—
69	Mount Apatite— Over exposed	33.0	—	—	—	—	—	—	—	—	—
31	Embryonic	37.0	31.2	—	25.2	—	—	—	16.3	—	—
56	"	36.2	31.3	28.8	25.2	22.2	21.0	—	17.0	—	—
64	Over exposed	36.6	—	—	—	—	—	—	—	—	—
63	Murray Bay— Over exposed	35.6	—	—	25.0	—	—	—	—	—	—
81	"	36.8	—	—	—	—	—	—	—	—	—
	Mean of first 20	34.4	30.5	27.1	23.1	20.6	19.7	16.9	15.6	12.4	9.5
	Average deviation %	0.7	2.3	2.8	1.8	1.7	1.9	2.0	3.0	3.5	0.8
	Air equivalent	6.97	6.19	5.50	4.69	4.18	4.00	3.43	3.16	2.52	1.93

78 by taking records of sections of the same haloes found on adjacent cleavage flakes. These records, Nos. 95 and 96 respectively, gave average radii. Five haloes had radii much greater than the mean. This may be due to large nuclei. Each of these haloes were so overdeveloped as to preclude the possibility of seeing even a large nucleus.

In the first 20 haloes measured (except No. 28) the nuclei were either too small to be measurable or were obscured by the intense blackening at the centre. Since there is good agreement among the measurements, it is concluded that the nuclei in most darkened haloes were also small. It is very difficult to calculate accurately the nuclear correction when needed. A few

$\alpha$ -particles emitted from the surface of the nucleus would penetrate furthest and the correction to be applied to them would be the nuclear radius. This correction decreases rapidly for  $\alpha$ -particles emitted from beneath the surface on account of the greater stopping power of the nucleus. Under our conditions of measurement we have assumed a correction of one-half the nuclear radius. This has been done only for No. 28, whose nuclear radius was measured to be about  $1\ \mu$ .

*Comparison with  $\alpha$ -Particle Ranges.*

In order to facilitate further discussion it is now necessary to introduce the integrated ionization curve for  $\alpha$ -particles. This is drawn in fig. 27 (a) for the uranium family in air at  $15^\circ\text{C}$ . and 760 mm. This curve is based on an average of the Bragg curves for RaC' obtained by Henderson\* and by Curie and Behounek.† These agree closely except at the beginning of the range, which portion is relatively unimportant. The ranges used in building up the curve are those given by Rutherford, Chadwick, and Ellis.‡ On fig. 27 is also shown an integrated ionization curve (b) for a source emitting nine  $\alpha$ -particles from each of the uranium series to one from each of the actinium series.

Many writers have stressed the striking correlation between certain features of these curves and the ring radii in the haloes. The most outstanding feature on our records is the outer ring, which may be ascribed to RaC'. Its mean radius is  $34.4\ \mu$ . Joly gives a mean value of  $32.8\ \mu$  for this radius while Kerr-Lawson obtains  $33.0\ \mu$  for the same kind of Murray Bay biotite. The large discrepancy between these results and the result obtained with the photometer can be attributed to the subjective errors inherent in visual observation. There is a tendency for the observer to set the line of his micrometer eyepiece an appreciable distance short of the RaC' range (as at R in fig. 26), owing to the inability of the eye to judge correctly the extreme limit of the blackening. We confirmed this discrepancy by visual observation of a few haloes of average diameter with a micrometer eyepiece. Our arrangements were not as suitable as those of Joly and Kerr-Lawson since our microscope did not have a revolving stage, which they found advantageous. Our results,  $31.5 \pm 0.6\ \mu$ , were even lower than theirs.

From the radius of the RaC' ring, the stopping power of the biotite may be determined. The (extrapolated) range of RaC' is 6.97 cm. Hence the con-

\* 'Phil. Mag.,' vol. 42, p. 536 (1921).

† 'J. Phys. Rad.,' vol. 7, p. 125 (1926).

‡ "Radiations from Radioactive Substances," p. 86 (1930).

version factor (the distance in mica equivalent to 1 cm. air at 15° C. and 760 mm.) works out to be 0.000494. If anything, this factor may be a little too large, as the last traces of ionization at the foot of the Bragg curve are found about 1 mm. beyond the extrapolated range. On the photometer records we probably estimate the ring radius somewhere between the points corresponding to the extrapolated range and the foot of the Bragg curve. The uncertainty involved is unlikely to be more than 0.5%.

This factor may be determined theoretically from the composition of the biotite. From an analysis of the Murray Bay mica (density 3.116) Kerr-Lawson calculated the conversion factor to be 0.000498, based on air constants.

The results of a commercial analysis of the Mount Apatite biotite are as follows :—

	%
SiO <sub>2</sub> .....	34.4
Al <sub>2</sub> O <sub>3</sub> .....	22.1
Fe <sub>2</sub> O <sub>3</sub> .....	2.1
FeO .....	22.3
TiO <sub>2</sub> .....	0.8
MgO .....	1.9
K <sub>2</sub> O .....	9.4
H <sub>2</sub> O .....	3.9
	96.9

The density was found to be 3.18.

Assuming the remaining percentage to be made up of Na<sub>2</sub>O, Li<sub>2</sub>O, MnO, and F in the usual proportions for biotite, the conversion factor is calculated to be 0.000485.

Thus the calculated conversion factors are in better agreement with the value found from the photometer records than with the values obtained with the micrometer microscope (*e.g.*, Kerr-Lawson's 0.000473).

The actual shape of the photometer records differs from that of the integrated ionization curve in a way which will be discussed later. There are, however, certain features of the two curves which correspond. The point A has already been shown to agree with the end of the range for RaC', yielding a conversion factor which agrees within the limit of error with those calculated from the chemical constitution of the micas. If the experimentally found factor be applied to the remaining points B, C, etc., we obtain the equivalent distances



in air shown in the last line of Table I. The most accurately measurable points C, D, F, H, and I are marked on the integrated ionization curve (b), fig. 27. It will be seen that D falls very nearly at the foot of that feature of the curve which is due to RaA. The point F falls very nearly at the foot of the next feature, which is due jointly to radon and polonium since their ranges are so near together. Similarly H falls close to the foot of the next rise due jointly to radium, ionium, and uranium II.

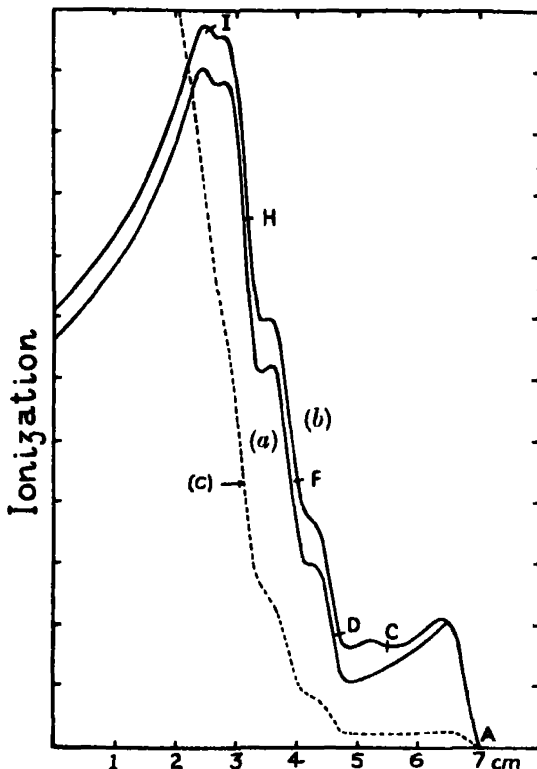


FIG. 27.

The integrated ionization curve given in fig. 27 differs from that given by previous writers (*e.g.* Holmes), in that it shows a double maximum at about 2.8 cm. This is caused by the use of newer and more accurate ranges for the two uraniums. It will be seen that the point I falls at the foot of the last rise in the curve, due to UI. This corresponds to the  $r_0$  ring recently discovered by Kerr-Lawson.

Thus all the  $\alpha$ -particles of the uranium family are accounted for. The remaining points B, E, G, and J are maxima. They also correspond to features on curve (b), fig. 27, but have been left off the curve to avoid confusion.

Furthermore, the point C occurs at the foot of a hump on the curve, which is due to the actinium family (An and AcC). This point will be discussed later.

This correspondence has often been pointed out by previous investigators. The present work is confirmatory and gives better agreement with air ranges on account of the greater accuracy of the halo photometer.

In the first three columns of Table II are given such previous measurements\* of external ring radii as can readily be identified with points on the photometer curves. The fourth column contains the mean values found in Table I.

Table II.—External Radii in Microns.

July.	Kerr-Lawson.	Yoshimura and Iimori.	Henderson and Bateson.
4th ring 32.8	R ring 33.0	32.8	A 34.4
3rd „ 21.7	$r_3$ „ 22.0	22.5	D 23.1
2nd „ 18.6	$r_2$ „ 18.6	20.6	F 19.5
—	$r_1$ „ 15.4	16.2	H 15.5
—	$r_0$ „ 12.3	—	I 12.7

### *Theory of Halo Formation.*

From the records of the halo photometer we have, for the first time, quantitative data on the blackening across a halo and on its modification in the various stages of development. The general features of the data are evident on inspection of figs. 3-17, Plates 12 and 13. It is now necessary to see if some theoretical explanation may be offered which will give numerical agreement with the records.

In drawing comparisons with the integrated ionization curve, all writers have had to point out that this curve is drawn for a bundle of parallel rays. In a halo the  $\alpha$ -particles radiate from a central nucleus and the law of inverse squares must be used in order to obtain the ionization per unit volume along a radius. This is done for the uranium family in fig. 27, dotted curve (c). It will be seen that the distinctive features of the curve for parallel rays have largely disappeared and comparisons with the haloes loses much of its significance.

In addition, some further assumption is needed to account for the various stages of development shown by the haloes. These haloes are so old that the

\* 'Nat. Res. Council Wash., Bull. No. 80,' p. 167.

whole uranium series must have long since reached equilibrium. The curves of fig. 27 are then the same for all haloes and unless some further factor is brought in, there can be no explanation of the infinite variety found among the haloes.

It is well known that many chemical reactions may be produced by ionizing agents such as  $\alpha$ -rays. It has been shown by Lind\* and others that many gaseous reactions are proportional to the amount of the  $\alpha$ -ray ionization producing them. This rule has also been extended with some success to the ionization due to other radiations such as  $\beta$ - and  $\gamma$ -rays. It has also been extended to reactions in liquids and solids by assuming that the total ionization produced by the complete absorption of radiation is the same as would be produced if the total absorption occurred in the same substance in the gaseous state.

The darkening of the mica is probably due to some chemical reaction, the exact nature of which need not concern us here. In attempting to use the integrated ionization curve as a measure of the extent of the reaction along a radius it must be remembered that this curve has been measured in air. There is no experimental method for finding the ionization produced along the path of the  $\alpha$ -particle in mica. We can, however, find the rate of loss of energy along the path in the mica.

Let  $E$  be the energy of the  $\alpha$ -particle at a distance  $x_m$  from the source in mica and let  $x_A$  be the distance measured along the range in air. The rate of loss of energy in mica is  $dE/dx_m = dE/dx_A \cdot dx_A/dx_m$ . Now Briggs† has measured very carefully the velocity of the particles from RaC' at different points  $x_A$  along their range in air. From these results he has deduced  $dE/dx_A$  which when plotted against  $x_A$  agrees fairly closely with the Bragg ionization curve for RaC'. In the same paper Briggs has also found the mass of mica per square centimetre equivalent in stopping power to 1 cm. of air at various points of the range. The reciprocal of this quantity is proportional to  $dx_A/dx_m$  and proves constant within the experimental error over the whole range of the  $\alpha$ -particle. From these results we conclude that the rate of loss of energy of the  $\alpha$ -particle in mica,  $dE/dx_m$ , is approximately proportional to the ionization in air as given by the Bragg curve. We then assume, as a first approximation, that the extent of the chemical reaction responsible for the blackening across the halo is proportional to the ordinate of the integrated ionization curve in air, corrected for the inverse square law, fig. 27.

\* "The Chemical Effects of  $\alpha$ -Particles and Electrons," 2nd ed.; New York (1928).

† 'Proc. Roy. Soc.' A, vol. 114, p. 341 (1927).

It is next necessary to consider the change in the system being acted on, namely, mica. To do this we have investigated the artificial blackening of the two micas by  $\alpha$ -particles. An apparatus was set up in which a narrow beam of light from an automobile lamp fell upon a photoelectric cell connected in series with a battery and galvanometer. A cleavage flake of the Mount Apatite specimen 5.5 mm.  $\times$  1.3 mm. in area and of thickness equivalent to about 5 cm. of air was placed in a thin glass capsule 1.3 mm. in diameter and 12 mm. long. The capsule was filled with about 29 millicuries of radon, sealed off, and fixed so that the beam of light passed normally through the mica on to the photoelectric cell. The current in the galvanometer was measured at intervals. The glass walls of the capsule were extremely thin and showed practically no discoloration in the course of the experiments. Consequently the observed decrease of the galvanometer readings was due to the blackening of the mica alone. The results are shown in fig. 28. The abscissæ represent the "exposures" on an arbitrary scale. By exposure is meant the total number of  $\alpha$ -particles emitted in the capsule from the moment of filling until the time at which the galvanometer reading was taken. The ordinates are galvanometer readings expressed in terms of the reading for unaffected mica as unity. When we use the word blackening in a quantitative sense we mean the galvanometer reading expressed in this way. It has not seemed worth while to complicate the discussion by introducing the density, since all experiments are made under comparable conditions.

A similar experiment was also carried out with the Murray Bay biotite. In this case the thickness of the flake was equivalent to about 1.3 cm. air and the quantity of radon was about 16.5 cm. The results are shown on fig. 28. The scales of exposure are not the same for the two micas. These curves are very similar to the blackening curve of a photographic plate, showing the features of under exposure, normal exposure, over exposure and reversal (for Murray Bay biotite). This would be shown clearly if the density were plotted against the logarithm of exposure, as is usually done for photographic plates. The phenomenon of reversal in mica exposed to  $\alpha$ -rays was first demonstrated by Jedrzejowski,\* who published a curve of the same general shape as that shown in fig. 28. It is not possible to compare accurately the exposure of the two biotites. It is clear, however, that the Murray Bay mica is much more readily blackened and shows clear signs of reversal, while the Mount Apatite mica may not have reached its maximum blackening with a considerably greater exposure.

\* 'C. R. Acad. Sci. Paris,' vol. 186, p. 135 (1928).

The results of the experiments on artificial blackening were applied as follows: some point was chosen arbitrarily as a reference point on a curve of type (c), fig. 27. In this case the point selected was the minimum at 4.8 cm. and the ratio of actinium to uranium was taken as  $1/20$ , corresponding more nearly to Mount Apatite mica. The ordinate at this point was then placed equal to the exposure at, say, the point P on the blackening curve, fig. 28. This could be done since both scales were arbitrary. The blackening of the halo at 4.8 cm. from the centre was then PQ. The blackening at any other point along the radius of the halo could then be calculated as follows. The

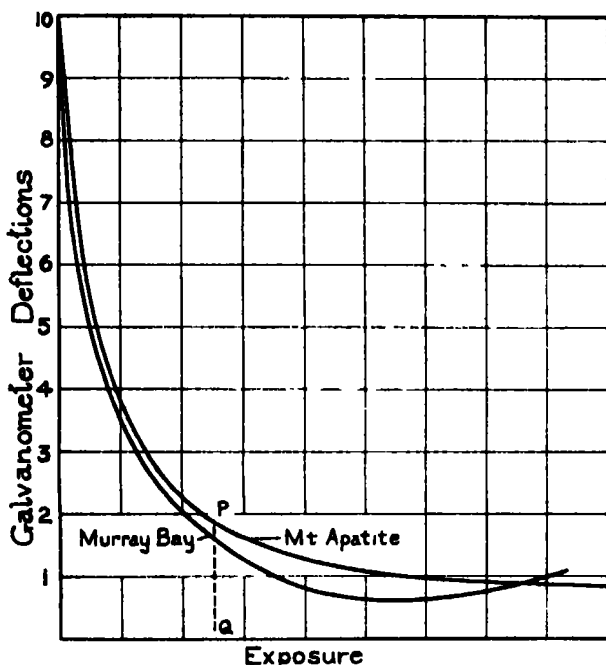


FIG. 28.

ionization at different points along the range on the integrated ionization curve was converted into the exposure (abscissa) on fig. 28 by the factor which had been fixed by choice of the point P. The blackening corresponding to this exposure was then read off the blackening curve, fig. 28. The results so obtained were plotted in fig. 29 (d), which should agree with the photometer record at some stage of development. Other stages could be represented by displacing the point P to other points on the blackening curve. A few of the family of curves so obtained are shown in fig. 29, a-d, arranged in increasing order of development. These curves have been drawn for the Mount Apatite mica.

The curves should be similar for the mica from Murray Bay, except for great exposure, when reversal would be evident at an earlier stage.

*Comparison with Experimental Records.*

By comparing the calculated curves on fig. 29 with the photometer records in Plates 12 and 13 it will be seen that there is fairly good general agreement.

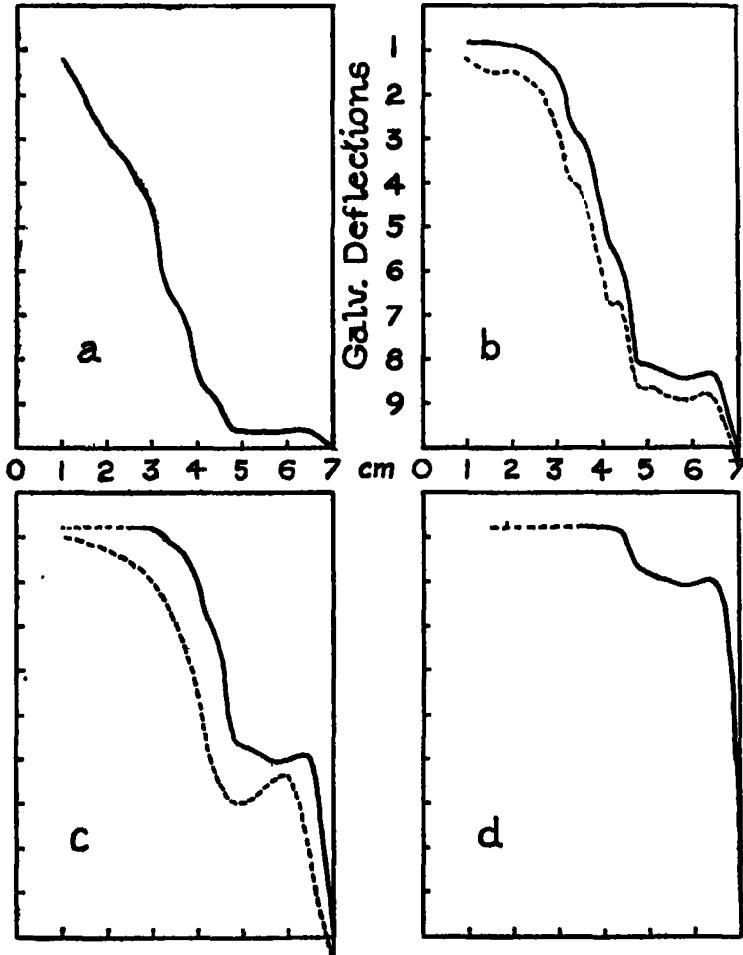


FIG. 29.

The gradual change of the family of curves with increasing exposure parallels very well the development of the halo as seen in the various stages. Two of the records which agree best are redrawn as the dotted curves *b* and *c*

respectively, fig. 29. These are displaced bodily downward a little to avoid confusion with the calculated curves.

There are certain points of disagreement between calculated and experimental results which should be mentioned. There are first the very striking dips in the photometer curve for a normal Murray Bay halo, which are associated with the decided ring structure of the haloes. From fig. 27 it will be seen that the integrated ionization curve (*c*), corrected for inverse squares, is nowhere convex to the axis of abscissæ except at the minimum at about 5 cm. Unless a minimum occurs on this curve it will not appear on the final curve calculated with the help of such blackening curves as are found in fig. 28. Secondly, there is disagreement for the haloes which we have called embryonic. This point will be discussed in the next section.

It may readily be shown that the addition of any percentage of the actinium or thorium families to the integrated ionization curve has the effect of smoothing out rather than accentuating the features which characterize that of the uranium family alone. Hence the disagreement cannot be assigned to this cause. The reason for the disagreement may lie in our use of the Bragg ionization curve as a measure of the agent bringing about the blackening. The Bragg curve has had to be used in default of better information. Unfortunately all measurements of the relationship between chemical reaction and ionization seem to have used the whole range of the  $\alpha$ -particles. It would be very desirable to have a measure of the relative efficiency of the ionization in promoting chemical reaction at different parts of the range. Such experiments are difficult with available radioactive sources. Just such a situation was present in our experiments on artificial blackening of mica, where the whole range of the particles had to be used in default of a sufficient intensity in a parallel bundle.

There is some reason to suspect that there would be some difference in efficiency at different points of the range, for the proportion of primary to secondary ionization varies markedly along the range.\* The accumulation of helium in the mica at the end of the range may also play a part in the blackening. It is, perhaps, significant that the Murray Bay mica, which is much more easily affected by the radiations, shows greater disagreement for the normal type of halo. To secure better agreement with experiment the Bragg curve should be modified so as to emphasize the maximum at the expense of the portion of shorter range. It is hoped that further information about the process of blackening may be obtained from a study of overlapping haloes.

\* Fowler, 'Proc. Camb. Phil. Soc.,' vol. 21, p. 531 (1923).

It cannot be claimed, of course, that the explanation of halo formation given above is complete, though it covers many of the observed facts. No attempt has been made, for instance, to explain what happens close to the nucleus, where some reversal should always be present if the theory were followed rigorously.

#### *Note on Embryonic Haloes.*

There is some doubt as to whether the haloes which we have called embryonic properly belong in the order of development to which they have been assigned above. It must be confessed that it is a little difficult to see how fig. 3, Plate 12, and fig. 10, Plate 13, fit into the general sequence of development shown in the rest of Plates 12 and 13, for the "shoulders" of these haloes seem to be too sharply accentuated. Moreover, such a shape in an early stage is not to be expected from the theory given above, as may be seen from fig. 29 (a).

The mean radius of the "shoulder" of a number of these haloes is  $16.3 \mu$ . Most of these contain visible nuclei whose mean radius is  $0.8 \mu$ . Subtracting a correction of half this amount gives the radius to the shoulder to be  $15.9 \pm 0.8 \mu$ . These haloes seem very similar to Joly's\* so-called "emanation haloes," the corrected mean of whose inner and outer radii is  $15.7 \mu$ .

The origin of these haloes must be considered as doubtful and further discussion is reserved until more evidence is available.

#### *Radioactivity of the Nucleus.*

No haloes attributable to thorium nor to  $\alpha$ -particles of unusually long ranges were found in the two biotites studied. There were some indications of  $\alpha$ -particles of unusually short range in the Murray Bay biotite, but further study is necessary before a definite value for the range can be assigned.

There was, however, a clear indication of the presence of the actinium family in the halo nucleus. A definite new ring, not hitherto reported, is found with a radius of  $27.1 \mu$ , equivalent to a range of 5.50 cm. From fig. 27 it will be seen that the radius of this ring C is found at the foot of the hump on the integrated ionization curve which is due to the addition of An and AcC. From the theory of the photometer curves given above, it may be shown that this ring should be noticeable only in normal haloes and that the effect of other members of the actinium series would be lost in the background of the uranium series. This was found to be true of both micas.

\* 'Phil. Trans.,' A, vol. 217, p. 51 (1917).



The hump may be clearly seen, for example, in fig. 12, Plate 13. It is interesting to note the appearance of this halo in the microscope, fig. 21, Plate 14. Under a magnification of about 350 the inner edge of this hump is just visible, but too faint to permit a good measurement with the micrometer eyepiece. It may be faintly seen on the original of fig. 21. (It is interesting to note that this hump is also to be recognized on the photometer records of the photo-micrographs published by Kerr-Lawson.) Thus the halo photometer not only gives a permanent quantitative record, but may bring out features of the halo not readily detectable by visual observation.

The significance of this actinium ring will be dealt with in another paper.

### *Conclusions.*

The qualitative picture of halo production given in the introduction is fully confirmed by the quantitative records of the halo photometer. The ranges of all the  $\alpha$ -particles, including uranium I, agree with the features of the halo with the greater accuracy obtained by the new method of measurement. The discoloration can be accounted for largely if not wholly by ionization. It is possible to predict quantitatively the main features and the progressive development of ancient haloes by calculations based on present-day experiments.

The halo photometer is well suited to the detection of unknown groups of  $\alpha$ -particles, particularly of very short or very long ranges. Such evidence would be of great interest and importance and a search for these groups has been begun. It is also intended to extend the work to thorium haloes and to haloes in materials other than biotite.

In conclusion, we wish to thank the National Research Council of Canada for a Bursary given to one of us.

### *Summary.*

A recording photometer has been constructed for the purpose of measuring the amount of darkening in pleochroic haloes.

Typical records of uranium haloes in biotite are given. They show features in good numerical agreement with  $\alpha$ -ray ranges, all the  $\alpha$ -particles from the uranium family being accounted for.

A theory of the formation of these haloes is proposed and found to be in general agreement with the photometer records.

The method serves to detect the presence of actinium along with the uranium and seems promising for the detection of possible  $\alpha$ -ray groups of hitherto unknown ranges.

---

## *A Quantitative Study of Pleochroic Haloes—II.*

By G. H. HENDERSON, Ph.D., King's College, and L. G. TURNBULL, M.A.,  
Dalhousie University, Halifax, N.S.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received January 29, 1934.)

[PLATE 15.]

The previous paper by Henderson and Bateson described a new method of studying pleochroic haloes with the aid of an instrument called a halo photometer. This method gave a quantitative permanent record of the structure of a halo, yielding accurate measurements of the various rings and also a record of the intensity of blackening throughout a halo. Thus an objective method took the place of the usual subjective study with the microscope.

The following paper gives further results obtained by this method and a discussion of the results.

### *Apparatus.*

Reference may be made to the previous paper for details of the apparatus. Here it is sufficient to say that an enlarged image of the halo, projected by a microscope, is moved slowly across a slit by a mechanism which is also geared to a rotating drum carrying photographic paper. The light passing through the slit falls on a photoelectric cell, the current from which passes through a galvanometer after amplification. The spot of light from the galvanometer is focussed on the paper on the rotating drum and a record of the variations in blackening across a halo diameter is obtained automatically.

The photocell previously used has been replaced by a vacuum cell (General Electric PJ 14), which had a greater resistance in the dark. The amplifier has been replaced by a simpler one using a single UX 232 screen grid tube, for which we are much indebted to Mr. H. A. Weir, M.A.

The enlargement factor of the instrument was found as before by taking frequent records of a test slide on which were engraved lines 0.01 mm. apart. The calibration was somewhat improved by making a careful survey of several revolutions of the screw with the help of the test slide. As the pitch of the screw was 1 mm. and the halo diameter about 0.07 mm., less than one-tenth of a revolution was used in making a record of the halo. The most uniform portion of the screw, as revealed by the survey, was always used in taking the records. Twenty-eight records of the test slide taken at intervals throughout

the work gave a mean of 2·048 cm. on the photographic paper equal to 0·01 mm. on the test slide with an average deviation of 0·017 cm. for individual records. On a single record the distances corresponding to the individual 0·01 mm. markings on the slide agreed to within 1%. Thus inaccuracies of the photometer mechanism should cause an error of less than 1% in the measurement of ring diameters.

### *Mica Specimens.*

We have examined a number of biotite specimens from various parts of the world. Fully half proved blank, while a number of others contained haloes too poor in quality for photometering. This was particularly true of the small flakes of biotite found in granites: work with these flakes is very difficult and time consuming. Micro-dissection with sharp tungsten needles is the only method we found suitable for obtaining flakes sufficiently flawless for use in the halo photometer. The flakes are often rich in haloes, but the nuclei are usually very large, causing over development, fuzzy edges and large and variable dimensions of the haloes. Our experience with these small flakes from the granites shows the need of great caution in interpreting these haloes. Ring diameters can be relied on only when the rings are sharp and the nuclei demonstrably small.

Naturally only a very small aggregate area of biotite flakes from the granite could be examined and very few suitable haloes were found. Those which were photometered gave records very similar to those discussed in the previous paper.

Larger crystals of biotite not only yield more perfect haloes but are more easily handled and better repay the time spent upon them. A specimen of light brown biotite from Egansville, Renfrew Co., Ontario, was found to contain very interesting haloes, and subsequent discussion in this paper will be restricted to this material. It was not particularly rich in haloes. Over two pounds of the specimen was split until the sheets covered an area of some 4 square metres, all of which was examined carefully under the microscope.

Nearly all the haloes were due to uranium (and associated actinium), but an occasional one contained thorium as well. In the earlier stages of development they were very similar to those discussed in the previous paper, but in the later stages their appearance was quite different. A series of bleached rings appeared in the interior of the halo, giving it a very striking appearance. This seems to occur in the "overdeveloped" stage. It is significant that in the Renfrew biotite we have never found haloes with uniformly blackened centres

and a paler outside ring, so characteristic of the overdeveloped stage in the Mount Apatite and Murray Bay biotites (*e.g.*, figs. 7, 15, and 22 of the previous paper). "Reversed" haloes are also found, in some of which the bleaching is very marked in the interior, only the outer edge of the halo being dark.

Photometer records of a few of these haloes are shown in figs. 1-8, Plate 15. They serve to give some idea of the variety and complexity found among the haloes from the Renfrew biotite.

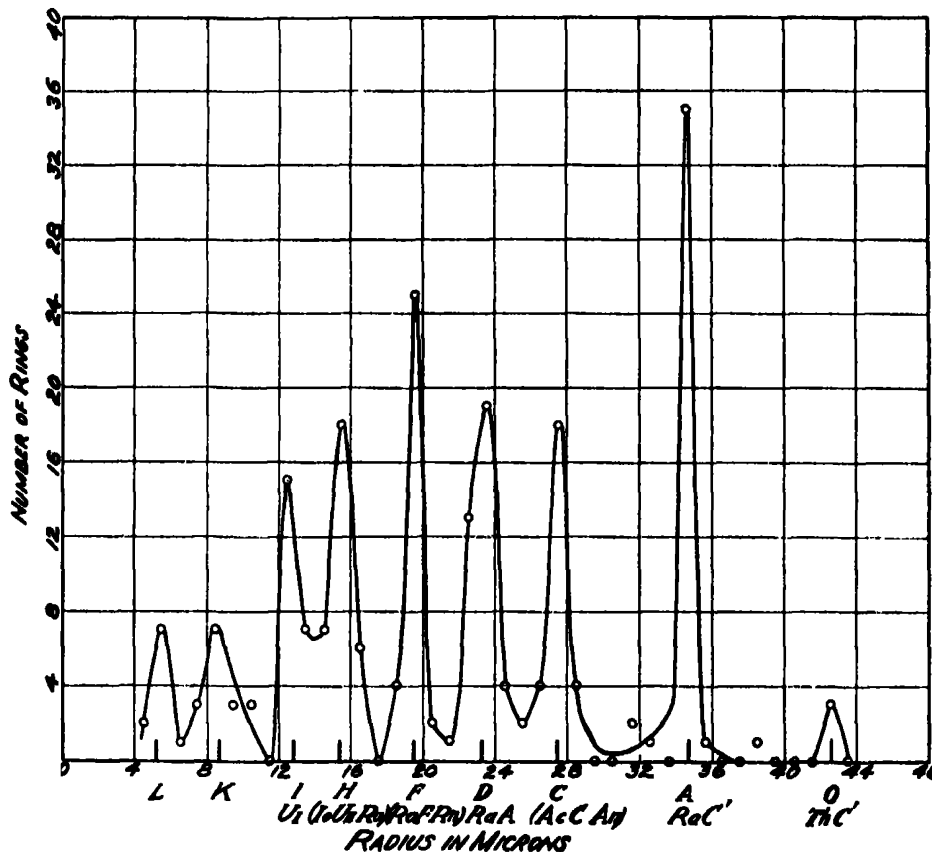


FIG. 9.

### Ring Radii.

The statistical distribution of the ring radii obtained from 40 haloes from the Renfrew biotite is shown in fig. 9. The abscissæ are the radii in microns, and the ordinates are the number of clearly defined rings found on the records, grouped in intervals of 1 micron.



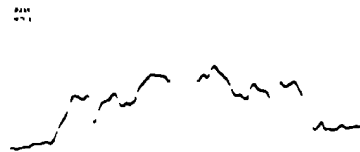
1



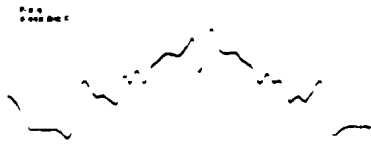
5



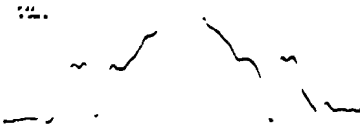
2



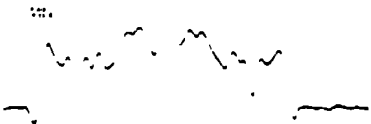
6



3



7



4



8



It will be seen that the rings fall definitely into well-defined groups. The mean of each group (from Table I) is shown at the bottom of the diagram and the associated  $\alpha$ -particle group is indicated below.

The mean values of the ring diameters are given in Table I. The letters in the first column follow the ring nomenclature previously used by Henderson and Bateson (fig. 26). The second column gives the mean radius in microns and the third the air equivalent in centimetres using the same conversion factor  $0.000494$  as that used in the previous paper. The fourth column gives the accepted range in air of the associated  $\alpha$ -particles. The fifth column gives the average percentage deviation of the individual rings from the mean.

Table I.

Ring.	Radius in $\mu$ .	Air equivalent cm. at $15^{\circ}$ C.	Associated $\alpha$ -particle.	% deviation from mean.
O	42.5	8.60	ThC' <sup>cm.</sup> 8.62	0.8
A	34.6	7.00	RaC' 6.97	0.6
C	27.4	5.55	{ An 5.79 AcC 5.51 }	1.9
D	23.2	4.70	RaA 4.72	1.7
F	19.4	3.93	{ Rn 4.12 RaF 3.92 }	2.0
H	15.4	3.12	{ Ra 3.39 U II 3.28 Io 3.10 }	2.6
I	12.8	2.59	U I 2.73	2.3
K	8.6	1.74	?	4.2
L	5.2	1.05	?	9.8

The ring radii used in determining the means given in Table I may be divided into two groups, viz., those for which a correction was applied for the finite size of the nucleus, and those needing no correction for this reason.

Twenty-one of the 40 haloes are subject to no such correction. They all showed good agreement among themselves and the nuclei, when visible, were small. Of these, 19 showed the RaC' ring clearly, giving an average radius of  $34.6 \mu$  with an average deviation from the mean of  $0.2 \mu$  or  $0.6\%$ . On account of the good agreement among the individual values and also with the value  $34.4 \mu$  found by Henderson and Bateson it was assumed that  $34.6 \mu$

was the correct value of the RaC' ring in this mica and that the conversion factor 0.000494 might also be applied to this mica.

The dimensions of the remaining 19 haloes were considerably larger than those of the uncorrected group. The nuclei when visible were large. It is difficult to measure with accuracy the size of the nucleus either from the records or with the micrometer eyepiece. Even if the size were known, the calculation of the correction to be applied on account of this is difficult and uncertain. Accordingly the excess of the RaC' ring (A) over  $34.6 \mu$  was taken as a constant correction for the finite size of the nucleus and was subtracted from all ring radii of that particular halo. (In two specimens when the RaC' ring was absent the RaA ring D was used instead.) The largest correction thus subtracted was  $2.9 \mu$  and the average correction for the 19 haloes was  $1.4 \mu$ . It should also be mentioned that in measuring some of the haloes allowance must be made for the Eberhard effect, which will be referred to in a later section.

The average percentage deviation from the mean given in column 5 of Table I are for the uncorrected haloes only. If the haloes which had been corrected for finite size of the nucleus were included, the percentage deviations would have been considerably smaller; for instance, that for the L ring would have been reduced to 6.0%.

No difficulty was found in assigning the rings to the groups making up the various means given in Table I. When the assignment was made with the help of fig. 9 it was found that all rings lying within  $1 \mu$  of the mean were included in the group while those with radii differing by more than  $1 \mu$  were all excluded. A few rings did not fall into any group, as may be seen from fig. 9. These are probably not real, but to be ascribed to flaws in the mica.

### *Discussion of Ring Radii.*

The results in Table I are in very satisfactory agreement with those of Henderson and Bateson on other micas. The differences are of the order of 1% with no systematic variation. As before, the ring radii are in good agreement with the accepted ranges of the  $\alpha$ -particles.

This agreement is significant in view of the ages of the three biotites concerned. Of the two biotites previously studied that from Murray Bay is of Pre-Cambrian origin, while that from Mount Apatite is supposed to be Devonian. The Renfrew biotite is probably Pre-Cambrian. These estimates, based on geological evidence, are confirmed by the ages deduced from the halo photometer records which will be discussed in a subsequent paper. Furthermore,



the stopping powers of the two first named biotites have been calculated from their chemical composition and found to agree within 1% of the experimentally determined value. Consequently we seem justified in the conclusion that the ring radii are the same in these biotites of different ages and also that the ranges of the  $\alpha$ -particles have remained the same throughout geological time.

It is well known that Joly's\* measurements of some ring diameters showed a progressive increase of radius with increasing age. From this he drew the conclusion that age determinations of minerals based on the radioactive constants might be seriously in error. The evidence† of later workers has been rather against this view. The more accurate results of the present experiments with the halo photometer are definitely against Joly's conclusion. A study of our measurements on the Renfrew biotite shows that either one of the rings H and I (due to U II, etc., and U I respectively) often occurs without the other in different stages of halo development and that one might easily be confused with the other in visual examination. It is to this cause that we are inclined to attribute the anomalous results found by Joly.

Apart from the rings due to the accepted members of the uranium and actinium families, three other groups appear in fig. 9. A very few of the haloes show a ring (marked O) at about  $42.6 \mu$ . This is undoubtedly due to ThC'. These rings were not found separately but associated with the uranium system in a compound halo. The intensity of the ThC' ring is so small that no clear indication of rings due to the shorter  $\alpha$ -particles of the thorium family is to be expected, nor were any found. An example of a record of one of these compound haloes is shown in fig. 8, Plate 15. No rings of radius greater than that of ThC' were observed.

Two groups of very small rings, marked K and L, are clearly discernible, corresponding to  $\alpha$ -particle ranges of 1.76 and 1.05 cm. respectively. Since they occur superposed on a dense background these rings are less clearly defined on the records. The accuracy of measurement is therefore less and a few of them must be classed as doubtful. Slightly more than half have had to be corrected for finite size of the nucleus. A record showing the K ring is shown in fig. 4, Plate 15.

Small rings have previously been reported. Joly's (*loc. cit.*) "hibernium" haloes have a radius  $5.0$ – $5.2 \mu$ , while Iimori and Yoshimura‡ find in their

\* 'Proc. Roy. Soc.,' A, vol. 102, p. 682 (1923).

† See Holmes, "The Age of the Earth," 'Nat. Res. Council Wash. Bull No. 80,' pp. 176-179, for a summary of the evidence.

‡ 'Sci. Pap. Inst. Phys. and Chem. Res. Tokyo,' vol. 5, p. 11 (1926).

$Z_1$  haloes a ring at  $5.4 \mu$ . Thus the existence of a ring at about  $5.2 \mu$  ( $1.05$  cm. air equivalent) seems well established. It may, perhaps, be significant that Hevesy\* and co-workers have recently found a range of  $1.13 \pm 0.04$  cm. for the newly discovered  $\alpha$ -rays from samarium. Joly and the Japanese investigators also report somewhat larger rings which may correspond to the ring which we find at  $8.6 \mu$ . The values which they give for the radius vary from  $6.0$  to  $10.8 \mu$ , so that it is difficult to draw comparisons.

No traces of these small rings had been found previously in the Mount Apatite mica and only faint traces were discernible in that from Murray Bay. Even if the emitters of these short range particles were a common constituent of all uranium bearing nuclei, it is probable that only a few biotites would develop sufficient structure to show these small rings and even then a very careful selection would probably be needed to find just the right stage of development. Further work on these small rings is very desirable.

It should also be pointed out that the distribution of fig. 9 is not typical, even of the Renfrew biotite. The haloes have been carefully selected so as to show the greatest number of rings. With a random choice of well-defined haloes there would have been a much greater proportion of the rings due to  $RaC'$  and  $RaA$ . In the early stages of halo development the two outer rings only are to be seen, while if the development has gone too far the inner detail is wiped out by blackening and bleaching. It should also be noted that we have made no attempt to distinguish rings of radius less than about  $4 \mu$ , to avoid confusion with possible extensions of the halo nucleus.

### *The Blackening of the Haloes.*

A comparison of the records reproduced in Plate 15 with those of the previous paper shows that the Renfrew haloes have a much more pronounced ring structure, a greater extent of bleaching and generally greater variations in blackening. In the embryonic and normal stages of development the Renfrew haloes are quite similar to those from Murray Bay, but the overexposed stage is radically different. It is in this stage that the striking appearance of the Renfrew haloes is manifested.

In discussing the Renfrew records we would first draw attention to the marked dip often found in the curves outside the  $RaC'$  ring. Examples are shown in figs. 3 and 4. At these dips the mica is evidently more transparent than in the region further out which has been unaffected by the  $\alpha$ -rays. This

\* Hevesy, Pahl, and Hosemann, 'Z. Physik,' vol. 83, p. 43 (1933).

seems to be the exact analogy of an effect well known in photography. When the blackening of an exposed and developed plate is registered by a microphotometer a decrease in the blackening sometimes occurs just outside the edge of a dense portion of the negative. This is known as the Eberhard effect. The explanation seems to be that when a strongly exposed area of emulsion is being developed a large quantity of soluble bromide is set free. This diffuses outwards and slows up the development of surrounding regions which have been less strongly exposed. It is noticeable only with certain types of plates and developers and may be minimized by special technique in developing.

In measuring the RaC' ring the Eberhard effect must be allowed for. The diameter measured at the bottom of the dip is from 1 to 3% greater than the diameter at the estimated level of the undisturbed surroundings, which was taken as the RaC' ring diameter.

It was pointed out in the previous paper that mica under the action of  $\alpha$ -rays shows the same phenomena of under, normal and over exposure and of reversal as does a photographic plate under the action of light. This parallelism is the more remarkable considering the different nature of  $\alpha$ -rays and light and the enormous difference in exposure time, millions of years instead of a fraction of a second. The Eberhard effect is another striking example of this similarity and requires further explanation.

To account for the Eberhard effect in mica the assumption might be made that there was some reaction, which we will call secondary, which darkened the whole of the mica in some late stage of halo development. Some such slow change in the mica is not at all unlikely in the course of its geological history. The secondary reaction would be hindered just outside the halo by the diffusion of the products of  $\alpha$ -ray ionization thus giving rise to the observed dips. The time is of course long enough for sufficient diffusion to take place in the solid.

Some support for this suggestion was found in the course of experiments on the artificial darkening of the Renfrew biotite by  $\alpha$ -rays. Small pieces of this biotite were sealed with radon in thin walled glass tubes and the progressive blackening was measured in the way described in the previous paper. The Renfrew biotite darkened slowly, but the maximum diminution in transmitted light was only 25% as compared with more than 90% for the two micas previously measured under similar conditions. These figures agree approximately with the diminution in transmitted light near the centre of the haloes as shown by the photometer records of the three micas. It seems improbable that these striking Renfrew haloes can occur in mica with such little suscepti-

bility to  $\alpha$ -ray exposure without the intervention of some secondary effect giving rise to further variations in blackening in the interior of the halo.

It will be recalled that in the previous paper a general explanation of halo formation was attempted, based on the integrated ionization curve for the uranium family with allowance for the inverse square law, fig. 27, c. This curve was built up of the Bragg curves of the individual  $\alpha$ -particles and represented very closely the rate of loss of energy of the  $\alpha$ -particles emitted by the halo nucleus. It was assumed as a first approximation that this curve also represented the extent of chemical action to be expected along a radius of the halo. Using the artificial blackening curve it was then possible to calculate the curves which should be obtained with the halo photometer. The results were in good agreement with the actual records, particularly for the Mount Apatite mica. In the Murray Bay records there occurred certain minima which could be accounted for by assuming that the chemical action along a single  $\alpha$ -particle track is represented not by the Bragg curve but by one having a somewhat sharper peak near the end of the range.

The same general explanation will serve for the earlier stages of halo development in the Renfrew haloes such as figs. 1 and 2, which are typical of a large number of these haloes. In the over developed stage, however, complications arise. The Eberhard effect will certainly begin to make itself manifest in the inner parts of the halo and the possibility of the secondary blackening mentioned above must not be lost sight of. These effects are not very amenable to calculation. While no contradiction is indicated with the general theory of halo formation outlined above, it must be recognized that a completely satisfactory explanation is still lacking for the over exposed haloes which are such a striking feature of the Renfrew biotite.

In conclusion, we wish to extend our thanks to Mr. C. M. Mushkat, B.A., who has given us much valuable assistance. For samples of biotite generously sent us we wish to thank Professor E. Backlund and Dr. E. Wiman, of Upsala, Sweden; Dr. S. Iimori, of Tokyo, Japan; Mr. A. R. E. Walker, of Cape Town, South Africa; Mr. Hugh D. Chase, of Orono, Maine, and Dr. H. V. Ellsworth and Dr. D. F. Kidd, of Ottawa, Canada. We also wish to thank Professor G. V. Douglas and Dr. G. O. Langstroth for valued discussions.

### *Summary.*

A study of a biotite from Renfrew, Ontario, was made by the halo photometer. The ring radii were found to be in satisfactory agreement with the

known ranges of the  $\alpha$ -particles of the uranium and actinium families. In addition two small rings were found having radii corresponding to  $\alpha$ -particles of range 1.74 and 1.05 cm. respectively. It is pointed out that there is no evidence of any change in the ranges of the  $\alpha$ -particles in the course of geological time. The Renfrew haloes yielded some very striking photometer records, on which could be seen an effect analogous to the Eberhard effect in photography. A discussion is given of the shape of the photometer curves and their possible explanation.

---

### *A New Method of Determining the Age of Certain Minerals.*

By G. H. HENDERSON, Ph.D., King's College, Dalhousie University,  
Halifax, N.S.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received January 29, 1934.)

The problem of age determination of a mineral is of great interest both from a geological and physical standpoint. The most widely used method is a radioactive one, based on the lead-uranium ratio as found by chemical analysis of the mineral. There are possible sources of error in this method owing to contamination with common lead and to alteration of the lead. It is the purpose of the present paper to suggest a new method which, although not at present capable of great accuracy, seems to be free from some of these difficulties. It is based on the presence of actinium in pleochroic haloes as disclosed by the new halo photometer discussed in the two preceding papers.

The most commonly accepted theory of the origin of the actinium series of radioelements, due to Piccard\* and Rutherford† is that the parent of the series is actinouranium, AcU, an isotope of uranium. This theory has been worked out in detail by von Grosse‡ while Western and Ruark§ have further analysed the relevant data. Since AcU is found to have a shorter period than UI, the older the mineral the greater should be the ratio of the amount of AcD lead ( $\text{Pb}_{207}$ ) to the amount of RaG lead ( $\text{Pb}_{206}$ ).

The actinium ring or hump on the photometer records is undoubtedly due to the  $\alpha$ -particles from AcC and actinon. When geological periods of time are

\* 'Arch. Sciences,' vol. 44, p. 161 (1917).

† 'Nature,' vol. 123, p. 313 (1929).

‡ 'Phys. Rev.,' vol. 42, p. 565 (1932).

§ 'Phys. Rev.,' vol. 44, p. 675 (1933).

being considered, the number of atoms of AcC which have formed and decayed, each with emission of an  $\alpha$ -particle contributing to the formation of this hump, is equal to the number of atoms of AcD which have formed. A similar statement is true of RaC' and RaG. Hence if the ratio  $\frac{\text{AcC}}{\text{RaC}'} = \frac{\text{AcD}}{\text{RaG}}$  can be found from the records the age may be calculated as follows.

Let  $N_1$  and  $N_4$  be the numbers of atoms of UI and AcU now present per gram of uranium and let  $n_1$  and  $n_4$  be the numbers of RaC' or RaG and AcC or AcD respectively which have been formed in the course of the time  $t$ . Further, let  $\lambda_1$  and  $\lambda_4$  be the respective transformation constants.

Then

$$n_1 = N_1 (e^{\lambda_1 t} - 1)$$

and

$$n_4 = N_4 (e^{\lambda_4 t} - 1).$$

Also if  $B = \lambda_4 N_4 / \lambda_1 N_1$  be the present branching ratio, i.e., the ratio of the numbers of AcU and UI atoms disintegrating per second we have

$$\frac{n_4}{n_1} = \frac{\lambda_1 B (e^{\lambda_4 t} - 1)}{\lambda_4 (e^{\lambda_1 t} - 1)}.$$

Thus knowing  $n_4/n_1$ , and the constants,  $t$  can be found. The solution of the equation is most easily found from a graph connecting  $n_4/n_1$  with  $t$ , such as that given by von Grosse. From this curve it is evident that the method is capable of reasonable accuracy for minerals whose age is greater than about 400 million years. For times less than this the value of  $n_4/n_1$  approaches asymptotically the limiting value  $B$ , which must hold for very young minerals.

#### *Interpretation of the Photometer Records.*

In order to find the ratio  $\frac{\text{AcC}}{\text{RaC}'} = \frac{\text{AcD}}{\text{RaG}}$  from the records it is first necessary to calculate how the ionization along a radius of the halo is affected by a given percentage of actinium and then to compare this with the ionization deduced from the photometer records.

Fig. 1 shows how the addition of actinium affects the ionization. The lower curve shows the integrated ionization curve for the uranium family alone, allowance being made for the inverse square law. The upper curve is obtained when one  $\alpha$ -particle of each of the actinium family is emitted for nine of each of the uranium series ( $\text{AcC}/\text{RaC}' = 1/9$ ). It will be seen that the ionization

due to AcA (6.58 cm. range) adds itself to the peak of the ionization due to RaC' and is practically indistinguishable from it. The ionization due to AcC (5.51 cm.) and An (5.79 cm.) together form a noticeable hump on the curve. This hump is very often found on the halo photometer records, and its radius agrees very well with that calculated from the ranges of AcC and An. The ranges of the shorter members of the actinium family are too close to those of the members of the uranium family to be distinguished from them. The

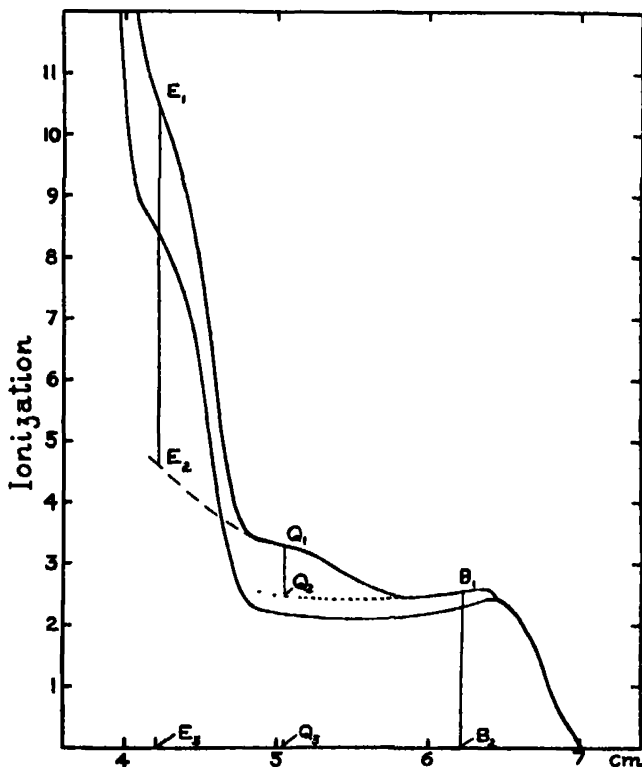


Fig. 1

— — — curve shows the ionization when the effects of RaA (together with RdAc and AcX) are deducted. The dotted curve shows the ionization to be expected when AcA alone is added to the uranium family. The points  $E_1$ ,  $Q_1$  and  $B_1$  have been placed on fig. 1 at the ranges corresponding to the top of the humps found on the photometer records.

If a given integrated ionization curve was known to be due to the uranium family plus an unknown percentage of the actinium family, this percentage could most easily be found by comparing the height of the An-AcC hump

above the background (formed by extrapolating the RaC' hump back towards the nucleus) with the height of the RaC' + AcA hump. On fig. 1 this extrapolation would follow the dotted line. It is found from fig. 1 that  $Q_1Q_2/B_1B_2$  is 0.32 when  $AcC/RaC' = 1/9$ . Thus the ratio of the An-AcC hump to the RaC'-AcA hump is 2.89 times the  $AcC/RaC'$  ratio. For other ratios it is found graphically that this factor varies only slightly, being 3.06 when  $AcC/RaC' = 1/20$  and 2.85 when  $AcC/RaC' = 1/7$ . Since these ratios of AcC to RaC' represent approximately the outside limits met with in practice, this factor may be taken as 3.0 with sufficient accuracy for present purposes.

A theory of halo formation has been proposed in the previous paper by Henderson and Bateson which was based on the integrated ionization curve for  $\alpha$ -particles and on artificial blackening experiments where the biotite is exposed to known numbers of  $\alpha$ -particles. Their theoretical curves for blackening in a halo were shown in fig. 29 and agreed fairly well with many of the actual photometer records. An inspection of these curves with special reference to actinium shows that a measurable effect will be found only for the An-AcC ring with the percentages of actinium found in practice. The effect of other members of the actinium family is lost when superposed on the background of the uranium family.

Furthermore, the effect is measurable with accuracy only in what was called the normal stage of halo development, since in embryonic haloes the effect is too small for accuracy while in the over exposed stage detail is largely wiped out.

The artificial blackening experiments for all biotites studied show that the blackening is proportional to the ionization in the early stages of the blackening. Thereafter the blackening increases less and less rapidly with increasing exposure to ionization. Hence for age determinations haloes should be chosen in the earliest stage of development consistent with accuracy of measurement. In such an early stage it should be possible to determine the  $AcC/RaC'$  ratio by comparing the humps on the photometer records in just the same manner as that outlined above for the ionization curve of fig. 1. Specifically, the ratio of the height of the actinium hump above its background to the height of the RaC' hump when divided by 3 should equal the required  $AcC/RaC'$  ratio.

Fortunately, there is a criterion which may be used in deciding whether the halo is still in the desired early stage of development. From fig. 1 it is seen that the intensity of ionization ( $E_1E_2$ ) of the hump due to RaA is about



twice the intensity of the RaC' hump ( $B_1B_2$ ). It may readily be shown from the artificial blackening curves that the blackening will be approximately proportional to the ionization occurring in the region between  $Q_1$  and  $B_1$  even when the RaA hump on the theoretical blackening curve of a halo\* is only one-quarter of that to be expected had the proportionality continued. It thus seems a reasonable criterion to include in the age determinations all photometer curves in which the ratio of  $E_1E_2/B_1B_2$  is greater than about 0.5.

### *Application of the Method.*

The method has been applied to three biotites, which have been studied in detail in the two preceding papers. For this purpose the curve given by von Grosse has been used, for which  $B$  is 0.04 and the period of AcU is  $4 \times 10^9$  years.

*Renfrew Biotite.*—Sixteen photometer records of haloes fairly free from flaws clearly showed the actinium hump and also satisfied the criterion mentioned above. The background of the actinium hump was drawn on each record by extrapolating the RaC' hump back towards the nucleus. The heights  $Q_1Q_2$  and  $B_1B_2$  were then measured. The following values were found for the ratio  $Q_1Q_2/B_1B_2$ : 0.21, 0.20, 0.28, 0.23, 0.34, 0.26, 0.20, 0.29, 0.14, 0.25, 0.23, 0.25, 0.20, 0.23, 0.23, 0.31. The mean is 0.240 and the average deviation is 0.037. The ratio  $AcC/RaC'$  is then  $0.240 \div 3 = 0.080$ . From von Grosse's curve it follows at once that the age of this biotite is 800 million years.†

*Murray Bay Biotite.*—Two photometer records were suitable giving  $Q_1Q_2/B_1B_2$  ratios of 0.203 and 0.254 with a mean of 0.228. Hence the  $AcC/RaC'$  ratio is  $0.228 \div 3 = 0.076$ , giving an age of 750 million years.‡

This is the same material as that used by Kerr-Lawson and is reported by Mawdsley to come from a pegmatite deposit cutting a Pre-Cambrian

\* Henderson and Bateson, *loc. cit.*, fig. 29.

† This biotite was secured as "pound material" from Ward's Natural Science Est., Rochester, N.Y., and is probably of Pre-Cambrian origin.

‡ [Note added in proof.—An analysis of uraninite from Murray Bay has recently been made by H. V. Ellsworth and F. F. Osborn giving  $Pb = 10.84$ ,  $U = 73.08$ ,  $Th = .088$  and hence an age of 1070 million years. It should be remembered that, as pointed out by Kerr-Lawson, the haloes are from secondary deposits and are not as old as the uraninite itself.]

*Mt. Apatite Biotite.*—Normal haloes are exceedingly rare in this material and only one record is suitable. The  $Q_1Q_2/B_1B_2$  ratio is 0.117 and the  $AcC/RaC'$  ratio is 0.039. All that can be said as to its age is that it is relatively young, probably less than 400 million years. It will be recalled that all young minerals give the same ratio 0.04.

I am indebted to Mr. Hugh D. Chase, of the Maine Geological Survey, for the information that the age is probably Devonian.

### *Discussion.*

Before going on to a critical discussion of the results a few remarks of a general nature may be interpolated here.

It is satisfactory that the ages of the three specimens are in agreement with the geological evidence. Moreover, according to the theory of the origin of the actinium series outlined at the beginning of this paper, haloes should not be found with rings due solely to the actinium series. No evidence of such haloes has been found in the many biotites examined. Thus the present experiments support the theory that the parent of the actinium family is an isotope of uranium.

It is also interesting to note that Nature has here given us a method of analysing quantitatively amounts of material far smaller than those used in any other method. The amount of uranium in a halo is of the order of  $10^{-13}$  gm. and the amounts of actinouranium and of other members of the two families are much smaller still.

It is difficult to estimate the accuracy of the age as given by the new method. The errors involved may be discussed under three heads, viz., those arising from the method itself, those arising from the measurements and those due to uncertainties in the actinium constants.

As regards the method itself, the theory of halo formation given in the two preceding papers has been found to reproduce the general shape of the photometer records. Special assumptions had to be made to explain detailed points, such as the minima on the records. It seems unlikely that such assumptions will give rise to any serious error in age determinations. Whatever be the mechanism responsible for the blackening in the halo, it will probably be the same for all  $\alpha$ -particles at the same distance from the end of their range. If the height of the actinium hump is measured above its background and if the criterion mentioned above shows that the outer part of the halo is in the early stages of blackening, this method of comparing the relative

number of  $\alpha$ -particles involved seems justified. A possible danger lies in the personal equation involved in extrapolating the background, but usually the photometer curves are such that there is only one reasonable extrapolation.

For the older biotites the errors arising from measurement of the photometer records are probably less than 10% when sufficient measurements are available. In this respect the results for Renfrew biotite are satisfactory, while those for the other two biotites serve more as illustrations of the application of the method. For very young biotites the method inherently gives only an upper limit for the age. The errors of measurement are large because the height of the actinium hump is small and the extrapolation of its background to some extent arbitrary. These errors are not due to the photometer itself since it has ample sensitivity and it repeats records of a halo in the same azimuth in a very satisfactory manner. They arise rather because of the difficulty of obtaining very flawless thin central sections of the halo. Small flaws and impurities introduce irregularities into the photometer curves, which may give rise to large errors of measurement and influence the extrapolation of the background. These errors are very much reduced by taking an average for a number of haloes. The problem involved may best be appreciated by reference to the actual records shown in the two preceding papers.

A more serious source of error lies in the present uncertainty as to the values of the period of AcU and of the branching ratio. It has been shown by von Grosse that the results of several detailed lead analyses are consistent with a period of  $\text{AcU} = 4 \times 10^8$  years and a branching ratio of 0.04. A study of the data by Western and Ruark indicates that the errors may be large. I am indebted to Dr. A. C. Lane, Chairman of the U.S. National Research Council's Committee on the Measurement of Geological Time, for calling my attention to the results of Marble and Aston for the lead from Great Bear Lake which indicate results inconsistent with von Grosse's values. Only a few leads have been fully worked out and further study is necessary for the elucidation of the actinium problem.

It is of interest to compare the method here proposed with the well-known method based on the lead-uranium ratio. It has been seen that the halo method requires an accurate knowledge of the actinium constants, which can only be found at present by the use of the Pb/U method itself. This method, however, is rather untrustworthy as it is subject to errors due to contamination and alteration of the radiogenic lead. (These errors are possibly also the cause of the present uncertainties in the actinium constants.) Progress in clearing up the resultant uncertainties in age determination has called for a

careful atomic weight determination and isotopic analysis of the lead. Such procedure is long and tedious and involves the co-operation of experts. While the addition of these refinements to the Pb/U method is practicable for a few selected mineral specimens it hardly lends itself to wide application. On the other hand, when the actinium constants are once known, the method here proposed is relatively simple, of wide applicability and appears to be free from the errors due to contamination and alteration mentioned above.

The halo photometer method seems capable of further development and of application to materials other than biotite. The work is being continued. I am anxious to obtain further samples of biotite containing haloes and would be greatly obliged to anyone who will send me specimens. Crystals larger than about 5 mm. are preferred.

#### *Summary.*

The effect of the actinium series in pleochroic haloes of the uranium type is discussed and shown to lead to a new method of finding the ages of the mineral bearing the haloes. Application of the method is made to three biotites which have been studied in detail.

---

*The Thermal Conductivity of Air by a Parallel Plate Method.*

By E. O. HERCUS, D.Sc., F.Inst.P., Associate Professor of Natural Philosophy, University of Melbourne, and D. M. SUTHERLAND, B.Sc., M.A., Bartlett Research Scholar, University of Melbourne.

(Communicated by T. H. Laby, F.R.S.—Received February 6, 1934.)

[PLATE 16.]

*Introduction.*

A measurement by a parallel plate method of the thermal conductivity of air was carried out by Hercus and Laby\* in 1913-14. The value then obtained, though higher than previous values, is considerably lower than that given by the more recent work by the hot-wire method of Weber,† Gregory and Archer,‡ and Kannuluik and Martin,§ and the parallel plate method differs fundamentally in principle from the hot-wire method. The latter, as has been pointed out,|| is subject to certain grave theoretical objections such as the difficulty of eliminating convection and at the same time allowing for the "Temperatursprung" at the wire. It is also open to question whether the geometrical measurements can be made with sufficient accuracy when a fine wire has to be placed along the axis of an ordinary glass tube. It was therefore considered desirable to repeat the measurement by the parallel plate method with improvements in thermocouple technique and improved mechanical construction. Measurements by this method are carried out at atmospheric pressure with complete absence of convection, and it is interesting to note that Weber, who was not aware of the previous work, suggested the desirability of carrying out the measurement by a parallel plate method.¶

*General Principle of Method.*

The general principle of the measurement is as described in the previous paper.\* Three horizontal copper plates, fig. 1, are separated from each other

\* 'Proc. Roy. Soc.,' A, vol. 95, p. 190 (1918).

† 'Ann. Physik,' vol. 82, p. 479 (1927).

‡ 'Proc. Roy. Soc.,' A, vol. 110, p. 91 (1926).

§ 'Proc. Roy. Soc.,' A, vol. 144, p. 496 (1934).

|| 'Phil. Mag.,' vol. 3, p. 1061 (1927).

¶ *Loc. cit.*, p. 484.

by small vertical distances. The central plate is in two separate portions, the central disc B being thermally insulated from the guard-ring D. If B, D, and the uppermost plate A are maintained at the same constant temperature  $\theta_2$ , while the lowest plate C is maintained at a constant lower temperature  $\theta_1$ , the heat supplied to the central disc B can escape only by conduction and radiation to D.

If  $E.I$  watts is the energy supplied to the central disc, then

$$EI/J = kS(\theta_2 - \theta_1)/x + \sigma S(\theta_2 - \theta_1), \quad (1)$$

where  $S$  cm.<sup>2</sup> is the area of B,  $x$  cm. its distance from D,  $k$ , cal. cm.<sup>-1</sup> deg.<sup>-1</sup> sec.<sup>-1</sup> the thermal conductivity of the gas,  $\sigma$  cal. cm.<sup>-2</sup> deg.<sup>-1</sup> sec.<sup>-1</sup>, the heat transfer by radiation, and  $J = 4.181$  joule per 20° C. calorie.\* In view of Stefan's law,  $\sigma$  is a function of temperature with a temperature coefficient in the neighbourhood of 20° C. of 0.0102 per ° C.

It was originally intended to determine the radiation either by (i) evacuating the apparatus to less than 0.0001 mm. Hg and measuring the heat transfer under these conditions, when air conduction would be negligible, or (ii) determining the heat transfer for a number of measured pressures from 0.01 mm. to 1 mm. Hg and extrapolating the linear (Knudsen) relation so found to zero pressure. Preliminary experiments disclosed serious difficulties in the use of either of these methods; consequently it was decided to eliminate radiation by measuring heat transfers at two plate distances. The transfer between the plates by radiation can be assumed to be the same at the two distances with like temperature conditions, for a calculation shows that, with the distances used, the radiation from the central plate to the cover through the narrow gap between the plates is not more than 1.4% of the total radiation. The effect of this on the final result is negligible.

If  $k_r$  is defined for a series of experiments at distance  $x_r$  by the relation :

$$EI/J = k_r S(\theta_2 - \theta_1)/x_r$$

then

$$k_r = k + \sigma x_r,$$

and  $\sigma$  can be eliminated if experiments are carried out at two distances  $x_1$  and  $x_2$ .

Corrections have to be made to equation (1) for heat transfers from B to A and D due to any slight inequality in temperatures of A, B, and D, and for the heat capacity of B should its temperature change during an experiment.

\* Laby and Hercus, 'Phil. Trans.,' A, vol. 227, p. 91 (1927).





FIG. 2. Photograph of upper half of central plate (The buttons shown as supports for the central disc were modified in the final design.)

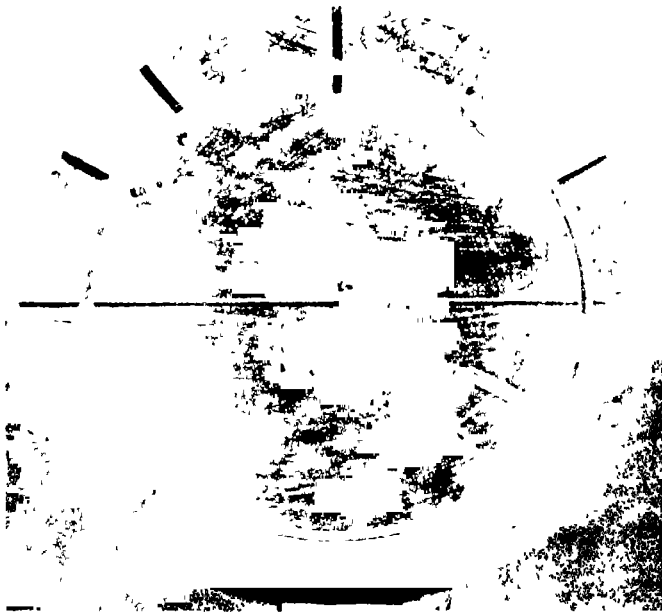


FIG. 3.—Photograph of lower half of central plate showing grooves for thermo-junctions

(*Face p. 691*)



*Apparatus.*

A diagram of the apparatus is shown in fig. 1. Each plate was made of two discs of copper each 6 mm. thick, soldered together after the necessary grooves for heater wires and thermocouples had been cut on their surfaces. Photographs of the two parts of plates B and D are shown in figs. 2 and 3, Plate 16.

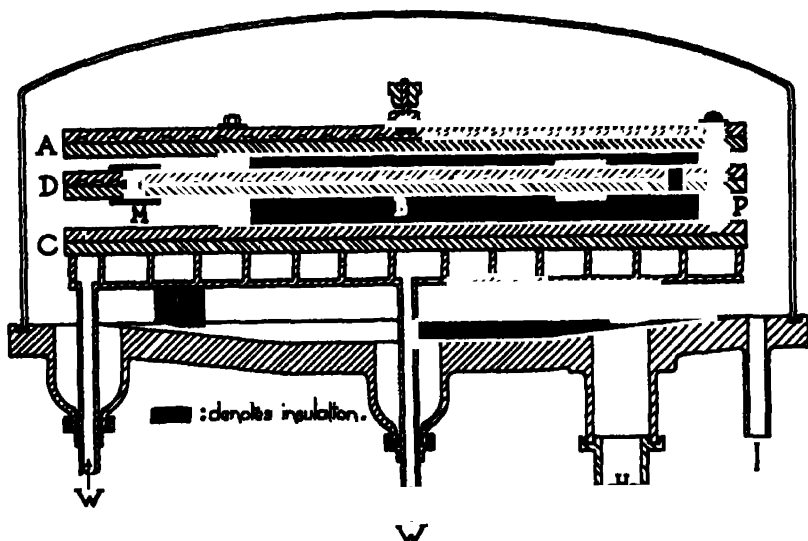


FIG. 1.—Diagram of conductivity apparatus. A, B, C are the three copper plates with guard-ring D. The centre plate B is supported from the guard-ring by three buttons, one of which is shown at M. The plates are separated from each other by two sets of three glass distance pieces shown at P. The water inlet and outlet are shown at W W. I is one of the 24 copper tubes through which the thermocouple and heater wires were led in, sealed by picein. H is the outlet to the vacuum pump and drying tubes. The spiral grooves for the heater wires are indicated, but not the radial grooves for the thermocouples.

As the gap ( $x_s$ ) between the central and bottom plates was only 6 mm. it was necessary for the two opposed copper surfaces to be extremely flat. The securing of this condition involved considerable time and difficulty, but was ultimately achieved by the mechanic staff of the laboratory. After the two halves of each plate had been soldered together with the wires inserted, the vital surfaces were tested with prussian blue on a surface plate and so scraped to within 0.001 inch of a true plane. They were then chromium plated under supervision and the surfacing process repeated. The result obtained is indicated in the end micrometer measurements given in Table I.

The heater grooves were cut as Archimedean spirals on the upper halves of the plates. On plate A provision was made for a terminal above the plate; on the centre plate B, the spiral was made double to avoid the need for a radial groove to carry the return lead. Radial grooves on the other halves were made to carry the thermo-junction wires. All wires were insulated by threading through small glass beads. It was found after the plates had been soldered together that this method of inserting the thermocouples was faulty in that they were directly affected by the heater wire; that is to say, the thermo-junction was hotter when the heater current was flowing than the plate whose temperature it was intended to measure.

This defect was eliminated by filling, in a vacuum, with molten paraffin wax, any voids which existed in the grooves between the wires inserted in them. The difference in temperature between the thermocouple and the copper plate surrounding it was investigated in the following manner:

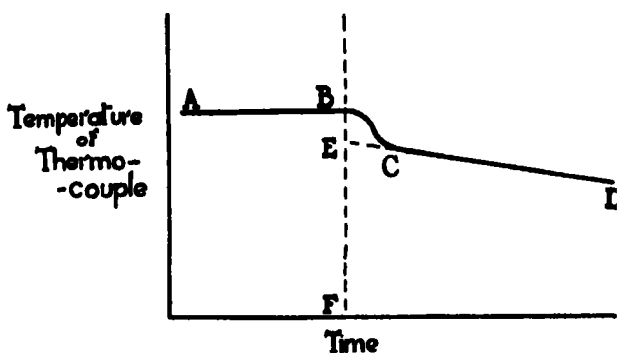


FIG. 4.

In fig. 4 AB represents the temperature of the thermocouple with the heater current flowing and adjusted so that the temperature is constant. At the time BF the current is switched off and BCD represents the thermocouple temperature for no heating current. CD is straight and is the cooling curve of copper plate; when CD is produced backwards it intersects BF at E. We take EB which was about  $0.02^{\circ}\text{C}$ . to represent the excess of the thermocouple temperature caused by the flow of heat directly from the heater wire to the thermo-junction.

On filling the grooves with paraffin wax the temperature difference BE became too small to be measurable.

Plate C formed the top of a brass box provided with a spiral partition for water circulation. Plates A, D, C were held together by three steel bolts, and separated by carefully ground glass distance pieces. Plate B was originally

supported from D by three ivory buttons, but these were found to be insufficiently rigid, and their place was taken in the final form of the apparatus by buttons of thin sheet steel, separated from the plates by mica washers to improve the thermal insulation. Unfortunately when these were made the plates had already been surfaced, so that it was not possible to let them in flush with the lower surface of B and D, and it was in consequence necessary to make a correction for the closer approach of their surfaces to the bottom plate C.

The whole plate system was mounted on a stout brass casting, supported on steel legs. A sheet steel cover could be waxed into a groove in this base. The thermocouple and heater leads were led in through picein seals in 24 short copper tubes, soldered into the brass base. Provision was made for connection to a vacuum pump, and for the water supply to the bottom plate. It was thus possible to pump out the whole system to a moderate vacuum and so secure a filling with air known to be dry and free from carbon dioxide.

To prevent the risk of any cross draught between plates B and C, a thin sheet of cellophane was attached round their edges, being provided with a small aperture between A and B.

### *Measurements.*

*Distance between Plates.*—This was determined by the thickness of the glass distance pieces, made in sets of three and ground parallel and of equal thickness. Two sets made were of thicknesses 0.2599 inch and 0.2343 inch respectively. The thickness was determined in terms of a set of Pitter block gauges by means of a fluid gauge, reading to 0.00001 inch. The pieces satisfied the requirements of equality and parallelism to at least 1 in 1000. Four maximum readings were taken for each distance piece, and the extreme variation over all readings for each set was 0.0002 inch.

Allowance had to be made, however, for the departure from flatness of the copper plates. This was done by temporarily substituting for the glass pieces a set of three equal steel ones of such a length as to enable measurements of the distances at selected points between plates B and C to be made with an end micrometer. At each of thirteen points distributed evenly over the bottom plate, end micrometer readings were taken of the distances between the plates. This "contour" for the final set of experiments gave the results shown in Table I.

The mean of the two observers is 3.1429<sub>5</sub> inch ; corrected for an error of 0.0004 inch in the micrometer it becomes 3.1425. As the mean length of the steel distance pieces was 3.1446  $\pm$  0.0002, the contour correction is -0.0021 inch.

Table I.—End Micrometer Readings.

Point.	Observer A.	Observer B.	Point.	Observer A.	Observer B.
<i>a</i>	3.1437	3.1435	<i>h</i>	3.1413	3.1418
<i>b</i>	33	33	<i>i</i>	399	03
<i>c</i>	15	19	<i>j</i>	421	23
<i>d</i>	24	26	<i>k</i>	433	38
<i>e</i>	37	40	<i>l</i>	433	41
<i>f</i>	44	46	<i>m</i>	429	32
<i>g</i>	40	43	Mean	3.1428	3.1431

### Area of Central Disc.

The diameter of the central disc was determined by a set of Pitter gauges and found to be 7.925 inches = 20.130 cm. The effective area (*S*) of the disc for conduction of heat is, however, slightly greater than this on account of the air gap between it and the guard ring. On the analogy of the similar problem as given by Maxwell\* for an air condenser

$$S = \frac{\pi}{2} (R^2 + R'^2) - \frac{\pi}{2} (R'^2 - R^2) \frac{\alpha}{x + \alpha},$$

where *R*, *R'* are the radii of central disc and inner edge of guard ring respectively, and

$$\alpha = \frac{R' - R}{\pi} \log_e \frac{2}{1 + e^{-\pi D/(R' - R)}},$$

where *D* cm. is the thickness of the plates.

Thus *S* varies slightly with *x* as shown in Table II.

Table II.

<i>x</i>	0.5898 cm.	0.8764 cm.	1.2499 cm.
<i>S</i>	331.0 cm. <sup>2</sup>	331.6 <sub>4</sub> cm. <sup>2</sup>	332.1 cm. <sup>2</sup>

\* "Electricity and Magnetism" (2nd ed.), vol. 1, pp. 280, 284.

*Temperature Measurements.*

All temperatures were measured by copper-constantan thermocouples, the latter being taken from one reel of 0.0076 inch diameter wire. A length of this wire was connected to a sensitive galvanometer, and a hot soldering iron was passed closely along it. A deflection of the galvanometer indicated an inhomogeneity, which was removed by cutting out a sufficient length of the wire. Even with these precautions it is difficult to ensure that no small inhomogeneities remain, and a subsequent kink in the constantan wire will develop one. Only continuous wires were used for thermocouples, as soldered joints in the wire were found to be the source of thermo-electric effects.

The various junctions were connected through a mercury cup selection switch with a Wolff three-dial "thermokraftfrei" potentiometer. The Kipp galvanometer used gave a deflection at about 2 metres of 5 mm. per microvolt.

*Calibration of Thermocouples.*

A standard couple was made up of wires from the same reels as the couples to be calibrated, and it was compared over the required range of temperature with a platinum thermometer, which had been calibrated in an investigation of the mechanical equivalent of heat\* so as to read accurately to 0.001° C., a well-insulated and stirred water bath being used for the comparison. This couple was then pressed into close contact with each of the copper plates in turn, and its readings noted against those of the couples embedded in the plates. Precautions were taken and tests made to ensure that the standard couple was actually at the temperature of the plate.

Ten readings of the standard couple when compared with the platinum thermometer showed that the following relation,† in which  $e$   $\mu$ V is the thermal e.m.f. and  $t^\circ$  C. the temperature, is correct to 0.2  $\mu$ V :

$$e = 72.96t - 13.892 (1 - e^{-0.00261t}).$$

The difference of  $e$  for the working couples from this relation, which at most was 5  $\mu$ V, was fitted to an accuracy of 0.5  $\mu$ V to a linear correction graph.

*Energy Measurement.*

The potentiometer used with the thermocouples was also used to measure the electrical input to the central disc heater, which was provided with current

\* Laby and Hercus, *loc. cit.*

† Adams, 'International Critical Tables,' vol. 1, p. 58.

and potential leads. In order to do this with one range of the potentiometer the shunt system shown in fig. 5 was used. The five resistances shown were standard coils made by Wolff or Hartmann and Braun, which, when compared with 100 ohm and 0.1 ohm coils standardized in 1933 by the National Physical Laboratory, were found to have an error of not more than 3 parts in 10,000. If the potentiometer readings are  $i$ ,  $e$ , microvolts respectively the current ( $I$ ) is  $0.0010101 i$  ampere and the potential ( $E$ ) is  $0.010001 e$  volts.

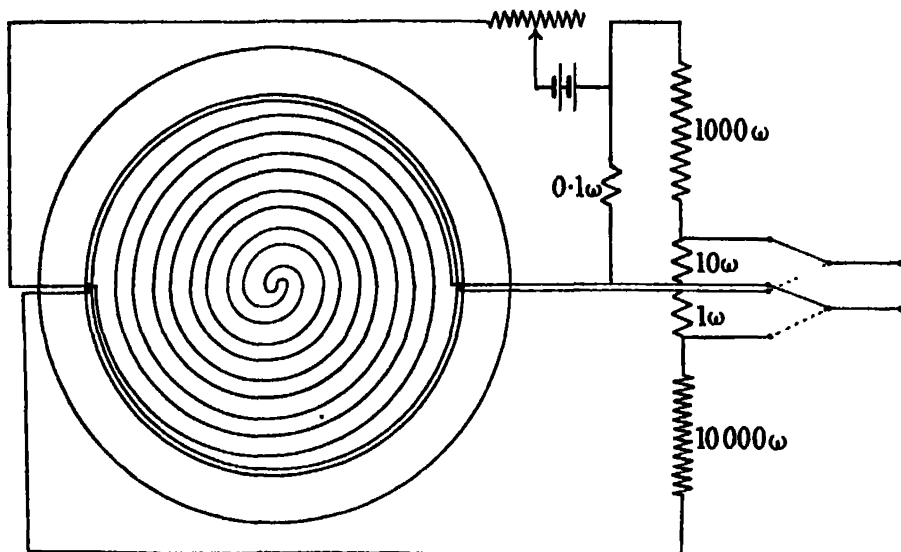


FIG. 5.—Circuit diagram showing shunts for power measurement by potentiometer.

The standard cell used throughout the experiments was compared with three cells certified in 1933 by the N.P.L. and found to be within 2 in 10,000 of its assumed value.

#### *Description of Individual Experiment.*

The general routine in carrying out an experiment was as follows: the water stream cooling the bottom plate was started, the heater currents in the other plates were switched on, and all plates brought as closely as possible to desired temperatures. A rough calculation was then made and the appropriate current for maintaining the temperature difference was passed through the centre plate heater. Checks were then kept of all plate temperatures, and adjustments made by trial and error until steady conditions were obtained. In a preliminary series of 45 experiments the conditions requisite for steadiness

had been investigated. Readings were then taken by one observer, and noted by the other ; generally for about 1 hour—as many readings as possible in that time being taken. Table III shows the readings for a typical experiment.

Table III.—Experiment No. 1. June 26, 1933.  $x = 1.2499$  cm.  
 $S = 332.1$  cm.<sup>2</sup>.

Time.		Centre plate heater.		Thermocouples ( $\mu$ V).				
		i.	e.	C.	B.	B — C.	D — B.	A — B.
h.	m.							
15	50.1	212.0	417.2	462.5				
	52.6				879.7	417.9	0.3	—0.2
	56.0	211.8	416.9	462.6				
	58.1				880.2	418.0	0.3	—0.1
16	01.0	211.7	416.9	462.8				
	03.0				880.3	417.9	0.4	—0.1
	06.0	211.7	416.8	462.9				
	07.3				880.6	417.9	0.5	—0.1
	11.0	211.7	417.0	463.2				
	12.7				881.0	418.0	0.3	—0.1
	16.0	211.9	417.1	462.4				
	17.5				881.2	418.5	0.2	—0.1
	20.0	211.8	417.0	462.7				
	22.3				881.7	418.7	0.2	—0.0
Means		211.8	417.0	462.7	880.7	418.1	0.3	—0.1

The columns headed C, B show the thermocouple readings for plates C and B respectively ; those headed B — C, D — B, A — B, the differential thermocouple readings for the plates indicated. Whence

$$I = 0.2139 \text{ ampere} \quad \theta_1 = 12.36^\circ \text{ C.}$$

$$E = 4.170 \text{ volt} \quad \theta_2 = 23.23^\circ \text{ C.}$$

Rise of centre plate temperature =  $2.0 \mu\text{V}$  in 29.7 minutes =  $0.0249 \times 10^{-3}^\circ \text{ C./sec.}$

Thermal capacity of centre plate = 256 cal./ $^\circ \text{C.}$

Therefore thermal capacity correction =  $0.02660$  watts and  $k_1' = 7.166 \times 10^{-3} \text{ cal. cm.}^{-1} \text{ deg.}^{-1} \text{ sec.}^{-1}$ .

This must be corrected for (1) heat gained from guard ring, and for heat lost to top plate ; (2) temperature coefficient of conduction (all values being corrected to  $20^\circ \text{C.}$ ) ; (3) " Button " error.

Thus

$$k_1' = 7.166 \times 10^{-5}$$

Gain from guard ring ..... 0.021

Loss to top plate..... -0.004

Correction to 20° radiation .. 0.022

Correction to 20° conduction 0.037

Button correction ..... -0.024

Therefore

$$k_1 = 7.218 \times 10^{-5} \text{ cal. cm.}^{-1} \text{ deg.}^{-1} \text{ sec.}^{-1}$$

### *Explanation of Corrections.*

(1) *Thermal Capacity Correction.*—The thermal capacity of the centre plate calculated from the dimensions of the plate, and the density and specific heat of copper is 256 calories per degree. It was also measured by eliminating  $k$  from the equations for two experiments; one with a rise, and the other with a fall, of the centre plate temperature. The value obtained was 261. The difference of 2% has a negligible effect upon the value of  $k$ , when the centre plate temperature alters only slightly during an experiment.

(2) *Loss to Guard Ring and Top Plate.*—Separate experiments using large differences of temperature between A and B, and D and B, were carried out. From them the heat transfers for given differences of temperature were calculated, and expressed as ratios of the heat transfer from centre plate to bottom plate. From these ratios corrections are made to the value  $k'$  calculated with top plate and guard ring losses neglected.

(3) *Temperature Coefficient of Radiation and Conduction.*—The mean temperature of the air layer is taken as the mean of centre plate and bottom plate temperatures. The temperature coefficient of radiation was calculated from Stefan's law; and is nearly constant over the small range of temperatures of our experiments. The relation,  $k_t = k_0 (1 + 0.00298t)$ , which is obtained by fitting\* a Sutherland formula to the observed values of  $k_t$ , is used to correct  $k_t$  for temperature.

(4) *"Button" Correction.*—The buttons supporting the centre plate projected 0.24 cm. from the plate, their surfaces being nearer to the bottom plate by that amount; this caused a slightly increased flow of heat from the area of the centre plate which they covered. Hence there is a correction to  $k$ , varying with  $x$ , which is calculated as follows: the temperature drop from centre

\* Laby and Nelson, 'International Critical Tables,' vol. 5, p. 214.



plate to bottom of the button is neglected since the button thickness is made up of 0.20 cm. of mica, and 0.04 cm. of steel, of conductivities roughly 30 times, and 2000 times that of air, respectively. Hence the whole button is taken to be at centre plate temperature. It is also assumed that half the heat conducted through the button comes from the centre plate, and half from the guard ring.

**Area of button face =  $3.14 \text{ cm}^2$ .**

**Therefore area of centre plate effectively covered by the three buttons**

$$= 3 \times 3.14/2 = 4.71 \text{ cm.}^2.$$

**Area of centre plate = 332 cm.<sup>2</sup>.**

Therefore fractional correction in  $k = \frac{4.71}{332} \left( 1 - \frac{x}{x - 0.24} \right)$

$$= 0.0034 \text{ for } x = 1.25 \text{ cm.}$$

0.0097      0.59 cm.

The results of the final series of experiments are shown in Table IV.

### Table IV.

No.	z cm.	Input watts.	Temperature in ° C.				k'.	k.
			$\theta_1$ .	$\theta_2$ .	$\theta_2 - \theta_1$ .	Mean.		
1	1.2499	0.86536	12.36	23.23	10.87	17.80	7.17	7.22
2		0.87310	12.39	23.29	10.90	17.84	7.21	7.22
3		1.2201	12.40	27.61	15.21	20.00	7.22	7.22
4		1.2118	12.43	27.60	15.17	20.01	7.22	7.23
5		1.0499	12.46	25.59	13.13	19.02	7.20	7.15
6		1.1916	12.41	27.27	14.86	19.84	7.22	7.20
7		1.1881	12.48	27.28	14.80	19.88	7.20	7.20
8		1.2038	11.97	26.92	14.95	18.46	7.25	7.25
9		1.3176	11.90	28.32	16.42	20.11	7.22	7.16
10		1.3257	11.93	28.34	16.41	20.13	7.27	7.21
						Mean $k_1 =$		7.21
11	0.5898	1.7388	13.30	24.56	11.26	18.93	6.57	6.57
12		1.7365	13.33	24.56	11.23	18.95	6.58	6.58
13		2.2820	12.30	26.96	14.66	19.68	6.65	6.60
14		2.1987	12.13	26.30	14.17	19.22	6.60	6.57
15		2.2711	11.95	26.59	14.64	19.27	6.60	6.59
16		2.2929	11.97	26.63	14.66	19.30	6.66	6.62
17		2.4086	12.13	27.44	15.31	19.79	6.70	6.65
18		2.4127	12.13	27.42	15.29	19.78	6.71	6.66
19		2.2054	12.30	26.43	14.13	19.37	6.64	6.60
20		2.1893	12.34	26.43	14.09	19.39	6.61	6.60
						Mean $k_2 =$		6.60

Therefore mean value  $k_{20}^{\circ} = 6.063 \times 10^{-5}$ .

Using temperature coefficient  $0.00298$  per  $^{\circ}\text{C}$ .,  $k_0 = 5.72 \times 10^{-5}$  cal. cm. $^{-1}$  deg. $^{-1}$  sec. $^{-1}$ .

The probable errors calculated from Table IV are, in  $k_1 \pm 0.018 \times 10^{-5}$ , in  $k_2 \pm 0.021 \times 10^{-5}$  giving the error of  $k_{20}$  as  $\pm 0.04 \times 10^{-5}$ , that is, 1 part in 143. As it is estimated that the length measurements are accurate to at least 1 in 1000, the power measurements to 1 in 2000, and the temperature calibrations to 1 in 800, the error in  $k_{20}$  is larger than is to be expected from these estimated errors. Two factors contribute to lower the degree of reproducibility of  $k_1$  and  $k_2$ , namely, (1) inhomogeneities in the thermocouples, and (2) small unsteadinesses in the temperatures in the plates and guard ring of the apparatus.

Before the series of experiments given in Table IV, 45 experiments were made in which the conditions for steadiness in the temperatures were investigated and considerably improved. The probable error for these preliminary experiments was  $\pm 0.08 \times 10^{-5}$ , twice that of the final series, showing, it is believed, that steadiness of conditions is the most important of the factors mentioned in attaining reproducible results. It is important to notice that there is no reason to suppose that such unsteadiness as remained would cause any *systematic* error in the final result.

#### *Comparison with the Hot Wire Method.*

The following are selected from the more reliable recent values obtained by the hot wire method :—

#### Thermal conductivity of air at $0^{\circ}\text{C}$ .

Gregory and Archer,* 1933 .....	$5.85 \times 10^{-5}$
Sophus Weber,† 1917 .....	5.68
Sophus Weber,‡ 1927 .....	5.74
Kannuliuk and Martin,§ 1933 .....	5.76

It will be seen that the present determination ( $5.72 \times 10^{-5}$ ), though carried out at atmospheric pressure and by a totally different method, agrees with the hot wire determinations, and that it agrees exactly with the mean of the last

\* 'Phil. Mag.', vol. 15, p. 301 (1933).

† 'Ann. Physik,' vol. 54, p. 325 (1917).

‡ *Loc. cit.*

§ *Loc. cit.*

three of the values given. In view of the greater ease of measurement by the hot-wire method it was decided to take the agreement for air as justifying the hot-wire method, and not to use the parallel plate method for other gases.

Our thanks are due to Professor Laby for encouragement and assistance throughout the work; also to the Australian Council for Scientific and Industrial Research, who provided the potentiometer and galvanometer used in the thermocouple measurements.

*Summary.*

The thermal conductivity of dry air at 20° C. at atmospheric pressure has been determined on the general lines of the method used by Hercus and Laby\* in 1914, but with an allowance for radiation by the use of two plate distances and with improved thermocouple technique and greater steadiness of conditions during the observations. The value found from a series of 10 experiments at each of two distances was  $6.06 \times 10^{-5}$  with a probable error of  $\pm 0.04 \times 10^{-5}$ . Using a temperature coefficient of 0.00298 per ° C., this reduces to a value at 0° C. of  $5.72 \times 10^{-5}$ , a result in good agreement with hot-wire determinations.

\* 'Proc. Roy. Soc.,' A, vol. 95, p. 190 (1918).

---

*The Probability of Inner Shell Ionization of Atoms by Electron Impact.*

By E. H. S. BURHOP, B.A., M.Sc., Dixon Research Scholar, University of Melbourne.

(Communicated by Professor T. H. Laby, F.R.S.—Received February 6, 1934.)

In the past few years a considerable amount of experimental work has been carried out on the investigation of the excitation of the optical energy levels of atoms by electron impact, and the probability of excitation of a given level has been measured as a function of the energy of the incident electrons. However, little has so far been done on the analogous problem in the production of characteristic X radiation, viz., the investigation of the relation between the probability of ionization of an atom in an inner shell by electron impact and the energy of the impinging electrons.

This problem has been discussed theoretically by Bethe\* using Born's approximate collision theory, which assumes that the impinging electron can be represented by a plane wave, and in which no account is taken of the distortion of this wave by the field of the atom with which it collides. The effect of this distortion will be negligible only when the energy of the incident electron is large compared with the ionization energy of the atomic shell with which it interacts. Under such conditions Bethe has found that  $f(V)$ , the probability of inner shell ionization of an atom by impact with an electron of energy  $Ve$ , can be expressed to a first approximation in the form

$$f(V) = \frac{k}{V} \log \frac{4V}{B_{nl}V_K}, \quad (1)$$

in which  $V_{Ke}$  is the ionization energy of the shell being investigated,  $B_{nl}$  is a parameter, constant for a given shell, and of the order of magnitude of unity and

$$k = \pi e^2 n^2 Z_{en} b_{nl} / R h Z_{en}^2, \quad (2)$$

where  $R$  is Rydberg's constant,  $Z_{en}$  the effective nuclear charge,  $Z_{nl}$  the number of electrons with principal quantum number  $n$  and azimuthal quantum number  $l$ , and  $b_{nl}$  a constant for a given shell. This formula would be expected to hold for values of  $V_i$  above about  $5V_K$ . The experimental investigation of

\* 'Ann. Physik,' vol. 5, p. 325 (1930).

$f(V)$  is thus of considerable interest in that it provides an independent check of the validity of the Born approximation.

The only direct experimental work on the problem appears to be that of Webster\* and his collaborators who have studied—for potentials up to seven times the excitation potential—the variation with tube potential of the intensity of  $K_{\alpha}$  radiation from thin silver targets less than 2000 Å. thick, the silver being in some cases in the form of a foil without any backing, and in other cases in the form of silver films deposited on beryllium. Such targets are so thin that the cathode rays used in the experiment were not appreciably diminished in energy in passing through them and thus the X-ray quanta were produced by practically homogeneous electrons.

The scarcity of experimental work in this field is due no doubt to the difficulties involved in obtaining thin targets which will not melt when bombarded with electrons of high energy. However, information on the form of  $f(V)$  can sometimes be obtained indirectly by using massive targets and measuring the variation with tube potential of the radiation emitted at very small angles to the target face. Thus for a target of solid silver a simple calculation shows that only 5% of the silver  $L_{\alpha}$  radiation emitted in a direction making an angle of 3 degrees with the target face comes from depths in the target greater than 3000 Å. Electrons of energy 30 kv. or more are not seriously retarded in passing through such a thickness of silver, so that for potentials above this, the form of the intensity potential curve will approximate to that given by a thin target. Measurements made under these conditions are described in this paper, and from them the form of the function  $f(V)$  for the  $L_{III}$  shell of silver is deduced in the range  $9V_K$ – $18V_K$ , a range in which the Born approximation should certainly be valid.

In addition a method of investigating the relative probability of ionization by electron impact of an atom in the K and L shells is described. Such an investigation may be used as a test of equation (2). A knowledge of this ratio is also of great importance in that it enables us to compare the relative intensities of the lines of the K and L series spectra—a comparison which is very difficult to make directly.

The method employed consisted in measuring the relative intensities of a number of L series lines—due to electron transitions ending on the  $L_I$ ,  $L_{II}$ ,  $L_{III}$  levels respectively—at two potentials, one about equal to and one about double the K shell excitation potential. It is clear that if an atom be ionized in a

\* 'Proc. Nat. Acad. Sci. Wash.,' vol. 14, p. 679 (1928); 'Phys. Rev.,' vol. 43, p. 839 (1933).

K shell, this shell will eventually be completed again at the expense of an electron from an outer shell and the emission of a K series quantum of radiation will result. Thus it would be expected that the K shell ionization of an atom would be followed by the ionization of the same atom in a shell further out (in the L shell if the K shell be completed with the radiation of a  $K_{\alpha}$  quantum, in the M shell if a  $K_{\beta}$  quantum be radiated, etc.). Such a result would also follow as a consequence of the Auger effect. In general, the relative "indirect" ionization of the  $L_I$ ,  $L_{II}$ ,  $L_{III}$  shells will differ from the relative ionization of the same shells produced directly by collision. Thus it is to be expected that the relative intensities of three L series lines arising from electron transitions ending respectively on the  $L_I$ ,  $L_{II}$ ,  $L_{III}$  levels will vary with the amount of K shell ionization. The magnitude of this variation provides a means of estimating the relative probability of ionization of an atom in the K and L shells by electron impact.

#### (I). *Experimental Details.*

*Description of Apparatus.*—The X-ray tube, with demountable anode, designed by Eddy,\* was evacuated with a four-stage Gaede mercury pump, and ran steadily at potentials up to 70 kv. with currents up to 5 milliamperes. A vacuum Bragg spectrometer was used, and was separated from the X-ray tube by a thin cellophane window. The calcite crystal could be rocked uniformly by means of a heart-shaped cam when making the relative intensity measurements. A valve-rectified constant potential generator, employing the Greinacher connection, served as a source of potential. The voltage output had a ripple of 3 kv. with a current of 5 mA. at all potentials, but, except at potentials below 20 kv., the correction for this ripple was negligible. The potential could be kept constant by adjusting the magnitude of the filament heating current and at the same time the tube current stayed constant to within a few hundredths of a milliampere.

The potential was measured by means of a Starke-Schroeder attracted disc electrostatic voltmeter calibrated by means of a standard spark gap and capable of reading the potentials in the required range to within 0.2 kv.

Current measurements were made with an Everett-Edgcombe precision milliammeter placed in the target end of the circuit so that it indicated, in addition to the tube current, the leak to earth through the anode water-cooling

\* 'J. Sci. Instr.' vol. 9, p. 354 (1932).

system. This leak was determined beforehand for various applied tube potentials. By using a telescope and mirror, current measurements could be obtained with this instrument to within 0.01 mA.

*Intensity Measurements.*—The intensities of the radiation investigated were recorded photographically, the films being photometred with a Moll microphotometer. Ilford double coated Rontgen film was used in these measurements and the films were developed for 4 minutes at 18° C. in the metol-hydroquinone developer recommended by the makers. All the films belonging to the one set of exposures were developed together, the films being developed with their surfaces vortical, and the solution kept agitated. A number of such sets of exposures were obtained for each line investigated, the potentials in each case being reduced to an arbitrary scale. In these experiments the line density was always kept less than 0.8 and it was assumed that the density of blackening of the film was proportional to the incident X-ray intensity. This has been verified over a wide range of wave-lengths for line densities less than 0.8 by several workers in this laboratory,\* using a technique similar to that used in these experiments.

*Preliminary Investigations.*—Before commencing the measurements an investigation was conducted of the behaviour of the tube at the different potentials. By taking pin-hole X-ray photographs of the target it was found that the position and size of the focal spot did not vary appreciably over the range of potentials used in the experiments. The position of the focal spot was checked from time to time during the measurements in the following way: the crystal was set for zero grazing angle of the X-ray beam, and behind it was placed a film which, on development showed the crystal cutting into the beam defined by the slits. Any change in the position of the focal spot would cause a change in the amount of the beam obscured by the crystal.

In these kinds of X-ray tubes the target sometimes becomes coated with tungsten evaporated from the hot filament, but usually this becomes serious only if the vacuum is insufficiently high. Photographs of the tungsten  $L_{\alpha}$  lines were taken from time to time during the measurements, but these lines were always found to be so faint that the effect of any deposition could be ignored. In an X-ray tube containing a number of wax joints it is also probable that carbon will deposit on the target face. It was not possible to test for this in the same way as for tungsten, but the target was removed and cleaned

\* Eddy and Laby, 'Proc. Roy. Soc.,' A, vol. 127, p. 20 (1930); Mohr, 'Proc. Roy. Soc.,' A, vol. 133, p. 292 (1931); Rogers, 'Proc. Phys. Soc.,' vol. 43, p. 59 (1931); Martin and Lang, 'Proc. Roy. Soc.,' A, vol. 137, p. 199 (1932).

after each set of exposures. Actually there never appeared to be any appreciable accumulation of carbon on it.

In reducing the observations it is assumed that the X-ray intensity is proportional to the tube current. This was verified independently by the author in a preliminary investigation for the range of current used in these experiments.

We shall first describe the measurement of the intensity-potential relation for the silver  $L_\alpha$  radiation.

(II). *The Thin Target Intensity-Potential Relation for Silver  $L_\alpha$  Radiation.*

*Theory.*—Consider a beam of electrons of  $V$  volt energy to be incident normally on the face of a target. In passing through the target the electrons will be diminished in velocity, the mean energy  $V_x$  of the stream at the depth  $x$  being given approximately by the Thomson-Whiddington law  $V^2 - V_x^2 = bx$ , where  $b$  is a constant characteristic of the target material. If  $i_0$  be the incident electron current, i.e., the number of electrons incident on the focal spot per unit time, then previous writers have written for the intensity  $I(V)$  of X radiation of the type being studied, leaving the target at an angle  $\theta$  to the surface,

$$I(V) = C \int_0^{\frac{V^2 - V_K^2}{b}} i_0 f(\sqrt{V^2 - bx}) \cdot e^{-\frac{\mu x}{\sin \theta}} dx, \quad (3)$$

where  $f(V)$  is the probability that an electron of energy  $V$  will produce ionization in an inner atomic shell,  $\mu$ ,  $V_K$  are respectively the absorption coefficient in the target material and excitation potential of the line being studied, and  $C$  a constant. Equation (1) is an integral equation which can be solved\* for  $f(V)$  in the form

$$f(V) = \frac{1}{Ci_0} \{bI'(V)/2V + \mu I(V)/\sin \theta\}. \quad (4)$$

Under the conditions of the experiment,  $\theta$  is small, and the second term in the bracket is much more important than the first, which thus occurs only as a small correction to the experimental curve.

The above treatment, however, is inadequate, as no allowance is made for the following four effects:—

- (1) The "diffusion" or multiple scattering of the electrons in passing through the target, as a result of which their length of path in the

\* This solution was given by Jönsson, 'Z. Physik,' vol. 43, p. 845 (1927).



effective 3000 A. of the target is increased, thus giving greater ionization than anticipated by the theory.

- (2) The "rediffusion" of the electrons due to the fact that some undergo large angle scattering near the surface of the target and re-emerge very little diminished in velocity. The proportion of the incident electrons that do this is surprisingly large. Schonland\* has found it to be 0.39 for a silver target, this factor being practically independent of the energy of the particles.
- (3) Owing to the rediffusion and absorption of the electrons in passing through the target material, the number of electrons per unit time crossing a section of the target parallel to the face diminishes rapidly as the depth of the section in the target increases.
- (4) Not all the measured radiation is excited directly by the impact of fast electrons. An indeterminate amount is excited by fluorescence arising from the absorption of general radiation.

The necessity for corrections (1) and (2) was pointed out by Wisshak† and they have been applied by Webster, Hansen, and Duveneck‡ to their latest results. These writers do not seem to have made any correction for the third factor, but the required correction would probably be very small for their results except at the lowest potentials used. The method of applying these corrections to the present results is a slight modification of that described by Webster, Clark, and Hansen.§

The theory of Bothell|| gives for the R.M.S. value of the deflection  $\lambda$  of an electron at a depth  $x$  in the target,

$$\lambda = \sqrt{2} \cdot \Theta, \quad (5)$$

where  $\Theta$ , the most probable deflection, is given by

$$\Theta = \frac{8}{\bar{V}} \cdot \frac{V + 511}{V + 1022} \cdot Z \left( \frac{\rho x}{A} \right)^{\frac{1}{2}}, \quad (6)$$

the potential  $V$  being measured in kilovolts,  $x$  in microns, the density  $\rho$  in gm./c.c.  $Z$  and  $A$  are respectively the atomic number and atomic weight of the target atoms.

For the treatment of the rediffusion correction a constant  $p$  is introduced which is the ratio of the number of electrons which rediffuse from a massive

\* 'Proc. Roy. Soc.,' A, vol. 108, p. 187 (1925).

† 'Ann. Physik,' vol. 5, p. 507 (1930).

‡ 'Phys. Rev.,' vol. 43, p. 839 (1933).

§ 'Phys. Rev.,' vol. 37, p. 115 (1931).

|| 'Handb. der Physik,' vol. 24, ch. 1.

target to the number incident. For a target of thickness  $x$ , Bothe gives for the same ratio,

$$p_x = p \cdot \frac{1 - e^{-2\alpha x}}{1 - p^2 e^{-2\alpha x}}, \quad (7)$$

where  $\alpha$  is Lenard's "practical absorption coefficient" for the incident cathode rays. Thus the fraction which rediffuses through the depth  $x$  under the conditions of the experiment is

$$p - p_x = p \{1 - (1 - e^{-2\alpha x})/(1 - p^2 e^{-2\alpha x})\}. \quad (8)$$

The velocity distribution of the rediffused rays from silver has been investigated by Wagner\* and by Chylinsky† who have found that in the range of potentials used in this work (30 to 60 kv.) their mean energy is about 0.8 of the incident energy. These rays also appear to obey approximately Lambert's cosine law of distribution in angle so that, as pointed out by Webster, rays rediffusing through a depth  $x$  of the target traverse on the average a distance  $2x$ . The rediffusion correction is very fully discussed by Webster, Clark, and Hansen in the paper previously cited.

Thus the intensity of the X radiation of the type studied, produced in the effective depth  $X$  of the target, instead of being given by the simple expression

$$C \int_0^X i_0 f(\sqrt{V^2 - bx}) \cdot e^{-\frac{\mu x}{\sin \theta}} dx \quad (9)$$

should actually be given by the equation

$$I(V) = C \left\{ \int_0^X i_0 e^{-\alpha x} \sec \lambda \cdot f(V) \cdot e^{-\frac{\mu x}{\sin \theta}} dx + \int_0^X 2i_0 (p - p_x) \cdot f(0.8V) \cdot e^{-\frac{\mu x}{\sin \theta}} dx \right\}, \quad (10)$$

together with a term representing the effect of the radiation excited by fluorescence after absorption of the continuous radiation.

Webster, Hansen, and Duveneck write this third term as

$$C' \int_{\nu_0}^{\nu} \int_0^{\pi/2} \{I(\nu_0, \nu)/\nu\} \{1 - e^{-\mu' x \sec \theta}\} \sin \theta \, d\theta \, d\nu, \quad (11)$$

where  $I(\nu_0, \nu) \, d\nu$  represents the intensity of the radiation in the continuous spectrum of frequency between  $\nu$  and  $\nu + d\nu$  when the maximum frequency excited is  $\nu_0$  and  $\mu'$  is the absorption coefficient of radiation of frequency  $\nu$  in the target material. The ratio of the constants  $C' : C$  is uncertain and,

\* 'Phys. Rev.,' vol. 35, p. 98 (1930).

† 'Phys. Rev.,' vol. 42, p. 393 (1932).

further, the form of the function\*  $I(\nu_0, \nu)$  is unknown for the values of  $\nu_0$  of this experiment. Most of the indirectly excited radiation is produced by frequencies not much greater than the frequency of the line being investigated, and a little consideration shows that if the laws of excitation of the continuous radiation are similar to those of the characteristic radiation, the form of intensity-potential curve for thin targets should not be greatly affected by making allowance for this means of excitation of the radiation. Experimental support for such a contention is supplied by the work of Webster,† and more recently of Hansen and Stoddard,‡ who have found that for massive targets of silver and palladium the amount of K series radiation produced by fluorescence bears to the amount produced directly by electron impact a ratio which is nearly constant over a wide range of potential. It was therefore decided to make no correction in this work for the fluorescent excitation of the radiation.§

*Experimental Observations.*—Fig. 1 shows the graph of intensity against potential for  $\text{AgL}_\alpha$  radiation, the angle of emission being approximately  $3^\circ$ .

Wave-length of  $\text{AgL}_\alpha$  radiation ..... 4.146 Å.

Excitation potential ..... 3.35 kv.

The corrections for diffusion, rediffusion, and absorption were found by integrating equations (1), (7), and (8) graphically using the following values of the appropriate constants:—

$$\begin{array}{lll} b|| = 4.5 \times 10^{12} \text{ volt}^2/\text{cm.} & \mu|| = 4850 \text{ cm.}^{-1} & X = 3000 \text{ Å.} \\ \alpha^{**} = 18,000 \text{ cm.}^{-1} & p|| = 0.39 & \theta = 3^\circ \ddagger \end{array}$$

\*  $I(\nu_0, \nu) = N(\nu_0 - \nu)$ ,  $N$  being a constant, for values of  $\nu_0$  near  $\nu$ , but the form of  $I(\nu_0, \nu)$  does not seem to have been investigated in the region important for this calculation, viz., for values of  $\nu_0$  many times that of  $\nu$ .

† ‘Proc. Nat. Acad. Sci. Wash.,’ vol. 14, p. 330 (1928).

‡ ‘Phys. Rev.,’ vol. 14, p. 339 (1928).

§ It might be thought that the unavoidable minute scratches in the target face would cause uncertainty in the value of  $\theta$ . To investigate this intensity-potential curves were determined for  $\text{AgL}_\alpha$ —

(1) with a target finished with 001 emery paper,

(2) with a target polished with rouge.

The intensity-potential curves obtained from the two targets were identical to within the limits of experimental error.

The target was always so turned that the final unavoidable minute scratches were placed parallel to the plane containing the axis of the tube and the direction of emission of the measured radiation.

|| Terrill, ‘Phys. Rev.,’ vol. 21, p. 476 (1923).

¶ Wien-Harms, ‘Handb. der Experimental Phys.,’ vol. 24, p. 234.

\*\* Andrade, ‘Structure of the Atom,’ p. 15.

†† Schonland, ‘Proc. Roy. Soc.,’ A, vol. 108, p. 187 (1925).

‡‡ The angle  $\theta$  could be measured to within about 20 minutes of arc. It is not necessary to know it more accurately.

In calculating these integrals the uncorrected experimental curve was taken to give an approximation to the form of  $f(V)$ . These calculations were made for potentials 35, 40, 50, and 60 kv. For potentials below 30 kv. the target ceases to approximate to a thin target. It was found that although the corrections to be applied to the experimental values to obtain values approximating to the conditions of the simple theory were large, they were

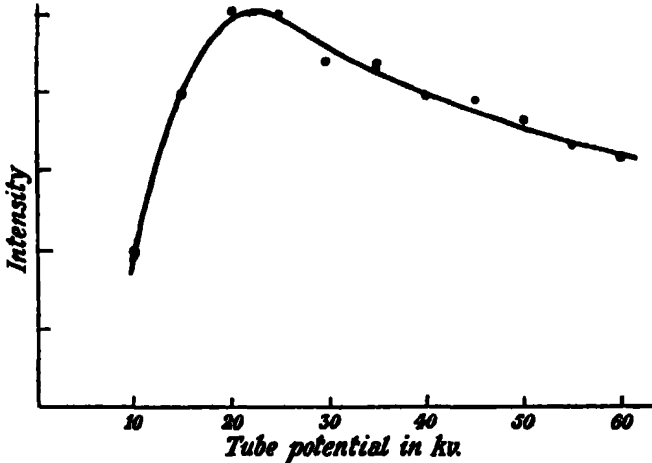


FIG. 1.—Intensity-potential curve for  $\text{AgLa}$ . (Angle of emission  $3^\circ$ .) (Each point is the mean of four determinations. Mean departure of individual determinations from curve = 3%.)

practically independent of the potential, so that the form of the curve was inappreciably altered by them.

Thus, denoting by  $I_0(V)$  the intensity of the radiation as determined by equation (9) and by  $I_{10}(V)$  the intensity as determined by the first two terms of equation (10), the table shows the ratio of  $I_0(V)/I_{10}(V)$  for different potentials.

	V = 35	40	50	60 kv.
$I_0(V)/I_{10}(V) =$	0.50	0.52	0.53	0.53
	(approx.)			

The value at 35 kv. is approximate only, as the method of calculating the diffusion correction is open to question at such low potentials. Applying equation (4) to determine  $f(V)$  it is found that  $\mu I(V)/\sin \theta$  is so large in comparison to  $bl'(V)/2V$  that the resultant form of  $f(V)$  differs inappreciably from that of  $I(V)$  in the range 30 kv. to 60 kv.

It is interesting to compare the results of the present investigation with those of Webster, Hansen, and Duveneck on the K shell of silver. These workers found their curves to be satisfactorily represented by Bethe's formula when the constant  $B_{ni}$  was given the value 6. Fig. 2 shows a comparison of the results of the present investigation with Bethe's formula ( $B_{ni} = 6$ ). Since the constant  $B_{ni}$  occurs only in a logarithmic term, the form of  $f(V)$  is insensitive to variation in its magnitude for the range of values of  $V/V_K$  of these experiments. It is seen that the curve deduced from the present work fits the theoretical curve to within 7 or 8% which is considered to be satisfactory. The curve obtained by Webster for silver K shell ionization is also shown on fig. 2 and the size of its ordinates is arbitrarily adjusted to fit on to the curve for silver  $L_{III}$  shell ionization obtained from the present paper to give the probable form of  $f(V)$  for inner shell ionization by electron impact in the range of  $V$  from 0-18  $V_K$ .

### (III). *Relative Probability of K and L Shell Ionization.*

Since the relative probability of K and L shell ionization will depend on the energy of the ionizing electrons it will first be necessary to define the conditions under which the determination is to be made. In this work the relative probability is measured as the ratio of the probability of K shell ionization of an atom by an electron of energy twice the K shell ionization energy to the probability of L shell ionization of the same atom by an electron of energy twice the L shell ionization energy.\* The ratio measured in this way will be said to be reduced to standard conditions.

*Theory.*—Let  $N_K$  and  $N_L$  be respectively the number of K and L shells ionized directly per unit time by the fast moving electrons.

Let  $I_{K_{a_1}}, I_{K_{a_2}}, \dots; \nu_{K_{a_1}}, \nu_{K_{a_2}}, \dots$ , be respectively the intensity and frequency of the  $K_{a_1}, K_{a_2}, \dots$ , lines of the element studied.

Let  $\phi$  be the ratio of the number of K quanta emitted to the total number of K shells ionized. ( $\phi$  in general is not unity owing to the existence of Auger's internal conversion effect.†)

Let  $\eta(V)$  be the ratio of the number of L shells ionized indirectly to the number ionized directly by the impact of electrons of energy  $V$ .

\* No distinction is made here between the different ionization energies of the  $L_I, L_{II}, L_{III}$  shells. As, however, this investigation only arrives at an upper limit for the relative K to L shell ionization probability, the error involved is not important.

† Auger, 'Ann. Physique,' vol. 6, p. 183 (1926).

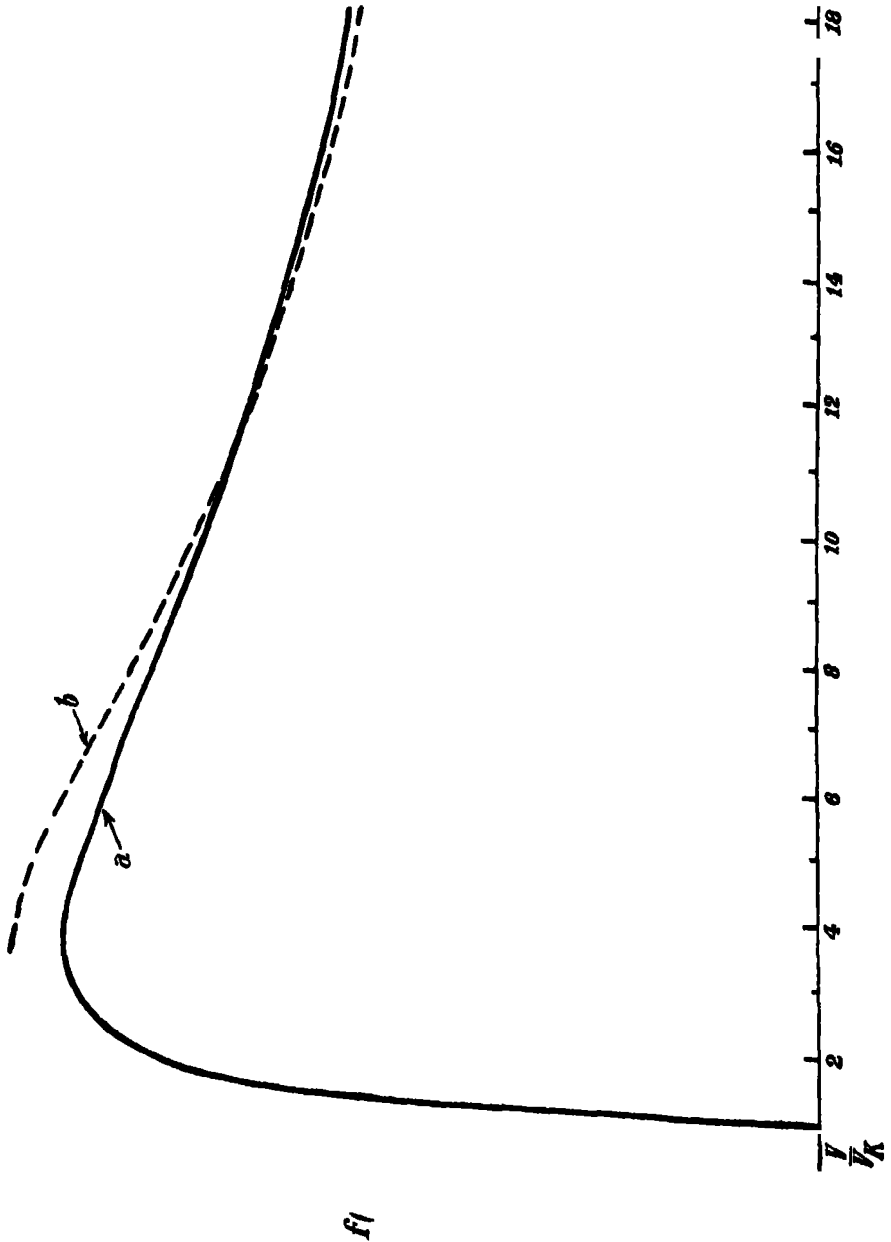


FIG. 2.

Curve *a* : Form of  $f(V)$ .

For  $V/V_K < 7$ , obtained from measurements of Webster, etc.

For  $V/V_K > 9$  obtained from measurements of author.

Curve *b* : Form of curve predicted by Bethe's formula (with  $B_{n1} = 6$ ).

Then, since the  $K_{a_1}$  and  $K_{a_2}$  lines arise when an electron switches from the  $L_{III}$  and  $L_{II}$  shells respectively to the K shell

$$\eta(V) = [1 - \phi + \{(I_{K_{a_1}}/\nu_{K_{a_1}} + I_{K_{a_2}}/\nu_{K_{a_2}})/(I_{K_{a_1}}/\nu_{K_{a_1}} + I_{K_{a_2}}/\nu_{K_{a_2}} + I_{K_{\beta_1}}/\nu_{K_{\beta_1}} + \dots)\}] N_K/N_L. \quad (12)$$

Equation (12) assumes that all the internally absorbed quanta are absorbed in the L shell. This approximation is justified since internal absorption produces only about 20% of the total effect in the case of silver, and the results of Auger suggest that most of the absorption occurs in the L shell.

Further let  $l_1 : l_2 : l_3$  be the ratio of the internal absorption coefficients of the K radiation in the three L shells.

Let  $r_1 : r_2 : r_3$  be the relative amounts of ionization of the three L levels due to electron impact; and  $\psi_{L_{a_1}}, \psi_{L_{a_2}}, \dots$ , etc., the ratio of the number of  $L_{a_1}, L_{a_2}, \dots$ , quanta emitted to the total  $L_{III}, L_{II}, \dots$ , etc., ionization.

Then at potentials below the K series excitation potential, the relative intensity of the  $L_{a_1}, L_{\beta_1}, L_{\beta_2}$  lines, involving respectively the  $L_{III}, L_{II}, L_I$  levels, is  $r_3\psi_{L_{a_1}}\nu_{L_{a_1}} : r_2\psi_{L_{\beta_1}}\nu_{L_{\beta_1}} : r_1\psi_{L_{\beta_2}}\nu_{L_{\beta_2}}$ , while above the K series excitation potential the relative intensity of the same lines is

$$\begin{aligned} & \psi_{L_{a_1}}\nu_{L_{a_1}}N_L \left( \frac{r_3}{r_1 + r_2 + r_3} \right) + \left[ (1 - \phi) \left( \frac{l_1}{l_1 + l_2 + l_3} \right) \right. \\ & \quad \left. + \frac{I_{K_{a_1}}/\nu_{K_{a_1}}}{I_{K_{a_1}}/\nu_{K_{a_1}} + I_{K_{a_2}}/\nu_{K_{a_2}} + I_{K_{\beta_1}}/\nu_{K_{\beta_1}} + \dots} \right] \psi_{L_{a_1}}\nu_{L_{a_1}}N \\ & : \psi_{L_{\beta_1}}\nu_{L_{\beta_1}}N_L \left( \frac{r_2}{r_1 + r_2 + r_3} \right) + \left[ (1 - \phi) \left( \frac{l_1}{l_1 + l_2 + l_3} \right) \right. \\ & \quad \left. + \frac{I_{K_{a_2}}/\nu_{K_{a_2}}}{I_{K_{a_1}}/\nu_{K_{a_1}} + I_{K_{a_2}}/\nu_{K_{a_2}} + I_{K_{\beta_1}}/\nu_{K_{\beta_1}} + \dots} \right] \psi_{L_{a_2}}\nu_{L_{a_2}}N_K \\ & : \psi_{L_{\beta_2}}\nu_{L_{\beta_2}}N_L \left( \frac{r_1}{r_1 + r_2 + r_3} \right) + (1 - \phi) \left( \frac{l_1}{l_1 + l_2 + l_3} \right) \psi_{L_{\beta_2}}\nu_{L_{\beta_2}}N_K. \end{aligned} \quad (13)$$

*Experimental Results.*—The results obtained are shown in Table I.

It is seen that the results of these measurements show no evidence of the predicted change of relative intensity of the lines investigated with potential for potentials above the K series excitation potential. Taking account of the accuracy of the measurements it is concluded that the magnitude of any such change of relative intensity must certainly be less than 5%.\*

\* These relative intensities have not been corrected for absorption in the cellophane window, etc. Such corrections are unnecessary for the present purpose. For the same reason it was not necessary to have a highly uniform sweep of the crystal so long as the characteristics of the sweep were reproduced in the different exposures.

Table I.

Lines compared.	Tube potential.	Relative intensity.	Number of readings.	Mean departure from the mean.
$L_{\alpha_1} : L\beta_1$	kv. 25	1.61	15	0.04
	55	1.57	9	0.06
$L\beta_2 : L\beta_1$	25	0.284	11	0.012
	50	0.281	13	0.013
$L\beta_3 : L\beta_2$	25	0.400	9	0.023
	50	0.409	10	0.009

*Interpretation of the Measurements.*—Before interpreting the above results in the manner suggested by the theory already given it is necessary to allow for some disturbing effects which would give rise to a small change in the relative intensity of the lines investigated.

Thus, the three L shells have slightly different excitation potentials. Assuming the function  $f(V)$  to have a similar form for all three L shells when  $V$  is expressed as a multiple of their respective excitation potentials, the ratio of the probability of ionization by electrons of energy 25 kv. to the probability of ionization by electrons of energy 50 kv. is not quite the same for the three shells, thus causing a change in the relative intensity of the L series lines at the two potentials.

Using the form of  $f(V)$  deduced from the first section of the paper, and taking the excitation potentials of the  $L_I$ ,  $L_{II}$ ,  $L_{III}$  shells of silver to be respectively 3.81 kv., 3.51 kv., and 3.35 kv., as deduced from the L absorption limit measurements given by Siegbahn, it is found that the change in relative intensity due to this cause should not exceed 1%.

Further, owing to the increased depth of penetration of the cathode rays at the higher potential a change in the relative intensity is to be expected owing to the difference in the target absorption of the lines. In these experiments the angle of emission of the radiation was kept small (approximately  $6^\circ$ ). Rough calculations using the simple formula (1) of the previous paper show that the change in relative intensity due to this factor is less than 2%.

Also, on account of the small angle of emission, no allowance is required for the retardation of the cathode rays in passing through the target material.

Taking into account these possible causes of variation in relative intensity and the probable error in the measured relative intensities, it appears that the relative intensities of the lines investigated are not different at the two potentials



by more than about 5% due to the filling up of the K shell and the internal conversion of the K series radiation.

In equations (12) and (13) have now to be put the values of the appropriate constants for silver. Thus :

For  $\phi$  Martin\* has given an empirical formula which predicts a value of 0.80 for silver ( $Z = 47$ ).

For the ratio  $I_{K_{\alpha_1}} : I_{K_{\alpha_2}} : I_{K_{\beta_1}} : I_{K_{\beta_2}}$  Siegbahn† gives 100 : 51.7 : 24.0 : 4.22 and for the ratio  $I_{L_{\alpha_1}} : I_{L_{\beta_1}} : I_{L_{\beta_2}} : I_{L_{\beta_3}}$ , 100 : 59.8 : 21.8 : 6. The values of  $\nu_{L_{\alpha_1}}$ , ..., etc., are taken from the same source.

Some evidence as to the relative magnitudes of the internal absorption coefficients,  $l_1$ ,  $l_2$ ,  $l_3$  of the K radiation in the L shell is obtained from the measurements of Robinson and Cassie‡ of the magnetic spectra of the photo-electrons produced by X-rays. For the elements they investigated it would appear that  $l_1$ ,  $l_2$ ,  $l_3$  are practically equal, and it has been assumed that this is so in these calculations.§

The constants  $r_1$ ,  $r_2$ ,  $r_3$  are assumed to be proportional to the number of electrons in the  $L_I$ ,  $L_{II}$ ,  $L_{III}$  shells respectively. Using the Pauli Exclusion Principle, Main-Smith|| predicts  $r_1 : r_2 : r_3$  to be 2 : 2 : 4.¶

Putting these values in equation (2), a short calculation shows that :—

- (1) The fact that the relative intensity of the  $L_{\alpha_1}$ ,  $L_{\beta_1}$  lines does not vary by more than 5% at the two potentials implies that the ratio  $N_K/N_L$  at 55 kv. must be less than 0.30.
- (2) The constancy of the relative intensity of the  $L_{\beta_1}$  and  $L_{\beta_2}$  lines implies that the ratio  $N_K/N_L$  at 50 kv. must be less than 0.30.
- (3) The constancy of the relative intensity of the  $L_{\beta_2}$ ,  $L_{\beta_3}$  lines implies that the ratio  $N_K/N_L$  at 50 kv. must be less than 0.05.

\* 'Proc. Roy. Soc.,' A, vol. 115, p. 420 (1927).

† "Spektroskopie der Röntgenstrahlen," Zwei Auflage.

‡ 'Proc. Roy. Soc.,' A, vol. 113, p. 282 (1926).

§ Actually Robinson and Cassie could not usually resolve magnetically the photo-electrons ejected from the  $L_{II}$  and  $L_{III}$  shells so that their measurements lead to the result  $l_1 : l_2 + l_3 = 1 : 2$ . Since the internal absorption produces only a small proportion of the total indirect effect, it is sufficiently accurate to assume in addition,  $l_2 = l_3$ .

|| 'Phil. Mag.,' vol. 50, p. 878 (1925).

¶ There appears to be doubt as to whether it is legitimate to separate the  $L_{II}$  and  $L_{III}$  shells in this way. For example, see Stoner, 'Proc. Leeds Phil. Soc.,' vol. 1, p. 226 (1928). The reductions were also made using for  $r_1$ ,  $r_2$ ,  $r_3$  the purely arbitrary values 2, 3, 3. Such a change had very little effect on the conclusions to be drawn from these measurements.

Applying now the form of  $f(V)$  deduced in the earlier part of the paper, one finds that if  $N_K/N_L$  at 50 kv. is less than 0.05, the ratio of the probability of ionization of silver atoms in the K and L shells by electrons of energy twice the respective ionization energies is less than 0.04.

One method of L shell ionization has been neglected in the above discussion, viz., the emission of a photo-electron from an L shell following the absorption of a quantum of continuous radiation. The difficulties involved in calculating this effect have already been pointed out. Its neglect can hardly affect seriously the conclusions to be drawn here as only an upper limit of the ratio of the K to L shell ionization probability is arrived at.

*Comparison with certain other Experimental Evidence.*—The only reliable evidence bearing on this work appears to be that of Pearson,\* who has investigated the relative intensity of the selenium  $K_{\alpha}$  and lead  $L_{\alpha}$  lines emitted from a target of lead selenide. His results lead to a value of about 0.4 for the ratio of the probability of selenium K shell ionization to the probability of lead L shell ionization, measured under the standard conditions. Allowing for the difference in atomic number of lead and selenium, and in the excitation potential of lead and selenium K series radiation, by the use of Bethe's formula it is found that the relative K to L shell ionization probability for lead, measured under the standard conditions is of the order of 0.01.

The results of one experimental investigation, however, are not in agreement with the above conclusions. Thus Stumpen† has investigated the relation between the intensity of tungsten L series radiation and the tube potential in a range of potential including the K series excitation potential of that element. He has reported an increase of slope of the intensity-potential curve when the K series excitation potential is exceeded. Interpreting his curves in the light of equation (12), it is found that they imply a value of the ratio of tungsten K shell ionization probability to tungsten L shell ionization probability, measured under the standard conditions, of about 0.5. His work is open to objection, however, on the ground that a pulsating source of potential was used. As the effect sought is small, it is necessary to know the applied potential accurately. It was therefore decided to repeat Stumpen's measurements. For this purpose a deep therapy Coolidge tube having a tungsten target was used, and an intensity-potential curve for tungsten  $L_{\alpha}$  radiation was obtained, using a technique similar to that described earlier.

Fig. 3 shows the curve obtained for an angle of emission of the radiation of  $45^\circ$ . No change of slope is evident at the K series excitation potential,

\* 'Proc. Nat. Acad. Sci. Wash.,' vol. 15, p. 656 (1929).

† 'Z. Physik,' vol. 36, p. 1 (1926).

although the curve has been determined over a range of potential wider than used by Stumpfen. This is consistent with a value of less than 0.04 for the ratio of tungsten K shell ionization probability to tungsten L shell ionization probability, measured under standard conditions.

The available experimental evidence thus appears to support the view that the ratio under discussion is of the order of magnitude of 0.01 and certainly

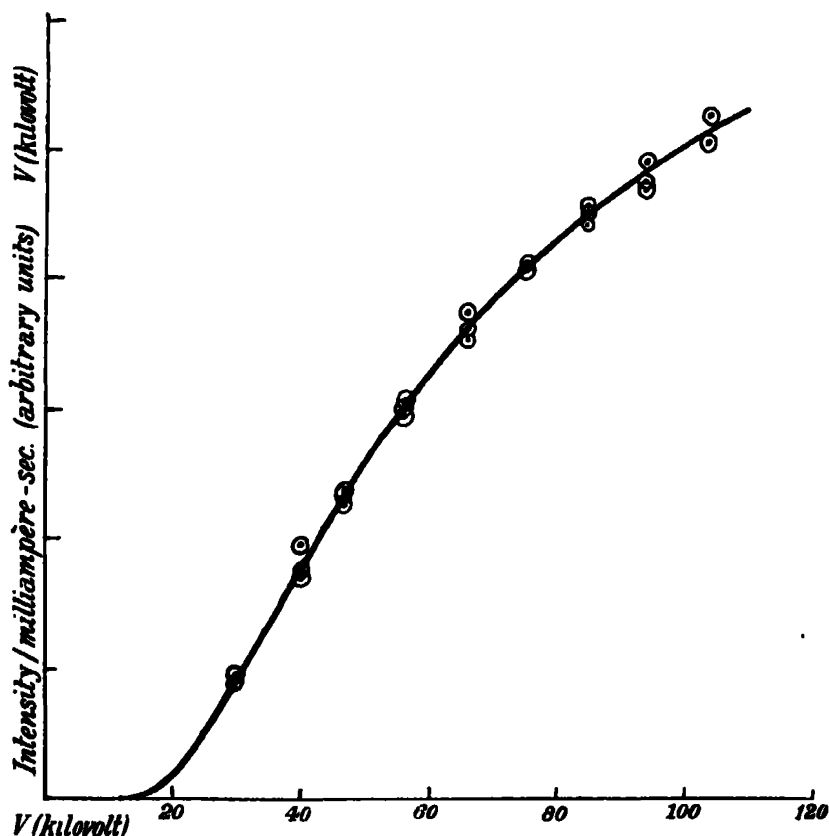


FIG. 3.—Thick target intensity-potential curve for  $WL_{\alpha}$ . (Angle of emission =  $45^{\circ}$ .)  
K series excitation potential of tungsten = 69.3 kv.

less than 0.04. It is interesting to compare this value with the result to be predicted from Bethe's expression for the constant  $k$  (equation (2)). Using the values of  $B_n$  calculated by Bethe, it is found that his formula predicts a value of the relative probability of slightly less than 0.01, in agreement with the order of magnitude expected from experiment.

**Conclusions.**—It is concluded that Bethe's formula, deduced from the application of the Born approximation to the problem of the inner shell ioniza-

tion of atoms by electron impact, predicts correctly the variation with incident electron energy of the probability of ionization of silver atoms in the  $L_{III}$  shell for the range of energies  $9V_K$ – $18V_K$ . The formula also gives correctly the manner in which the ionization probability varies from K shell to L shell.

*Acknowledgments.*—It is a pleasure to acknowledge my indebtedness to Professor T. H. Laby, F.R.S., who suggested this investigation, and to Drs. C. E. Eddy and L. H. Martin for their continued interest and encouragement, and for many useful discussions.

### *Summary.*

The probability of ionization of silver atoms in the  $L_{III}$  shell by electron impact has been measured as a function of the energy of the incident electrons for the range of energies  $9V_K$ – $18V_K$ ,  $V_K$  being the ionization potential of the shell. For this purpose intensity-potential curves for the silver  $L_\alpha$  line were obtained for potentials 15–65 kv. A massive target of silver was used and the rays which entered the spectrometer slit were emitted at a small grazing angle to the target face. In this way curves equivalent to thin target curves were obtained.

(2) It is predicted that the relative intensity of a number of L series lines should vary with potential for potentials above the K series excitation potential, and the magnitude of this variation should depend upon the relative probability of ionization of an atom by electron impact in the K and L shells. This is made use of in an attempt to measure this relative probability for silver atoms, and it is concluded that its value measured at twice the respective K and L shell ionization potentials is certainly less than 0.04.

The results of both parts of the investigation are in substantial agreement with Bethe's theory of inner shell ionization in which the Born approximation is applied to the discussion of the problem.

---

## *An Elementary Discussion of Ferromagnetism.*

By FRANCIS BITTER, John Simon Guggenheim Memorial Fellow.

(Communicated by R. H. Fowler, F.R.S.—Received February 7, 1934.)

The theory of ferromagnetism, as it is usually formulated, suffers from certain inconsistencies, especially as regards the definition of spontaneous magnetization. In this paper the situation is discussed from an elementary point of view, that is, considering only the statistical aspect of the problem, and omitting the more fundamental physical treatment of the molecular fields. It is intended to show as simply as possible, first, the origin of some of the undesired results of the Weiss Equation; second, the type of solution which a correct theory would probably give for perfect crystals; and third, to give a derivation of equations similar to the Weiss Equation, but from which it becomes apparent without further assumptions that the results cannot be applied indiscriminately to the whole magnetization curve. Finally an attempt is made to correlate the predictions of the theory with experimental data on the magnetization of actual crystals.

Fundamental to any discussion of ferromagnetism is the idea of spontaneous magnetization of Weiss, mathematically expressed in his famous equation, now slightly modified by the introduction of spacial quantization. Based on this we have the work of Akulov and Becker showing how various properties of crystals and elastically distorted materials may be described in terms of a few empirical constants, the number of these constants being limited by the structural symmetry of the material. Almost all the applications of Weiss's spontaneous magnetization to actual problems require the assumption that the material is not spontaneously magnetized in its entirety, but that only small regions are so magnetized, the resultant observed magnetization being determined by the distribution of the direction of magnetization of these regions. This assumption, however, is a flat contradiction of Weiss's Equation, which states that spontaneous magnetization exists irrespective of the size of the sample under consideration. In spite of the very considerable progress that has been made in recent years in our understanding of ferromagnetic phenomena the situation is not altogether satisfactory, primarily because these regions of spontaneous magnetization are introduced into a theory which specifically denies their existence.

The usual derivation of the Weiss equation is the following. By the usual methods of statistical mechanics the Langevin formula for the magnetization

of a paramagnetic gas is derived, no interaction between the elementary magnets being assumed. This gives

$$\frac{I}{I_0} = \coth x - \frac{1}{x}$$

$$x = \frac{\mu H}{kT}$$
(1)

the symbols having their usual significance. By substituting for  $H$  the expression

$$H = H_e + NI,$$
(2)

$H_e$  being the externally applied field and  $NI$  an internal field proportional to the existing intensity of magnetization, we arrive at the desired equation. Equation (1) becomes

$$\frac{I}{I_0} = \tanh x,$$

if we assume that the elementary magnets can point only parallel or antiparallel to the field. From such considerations we can obtain the energy relations for ferromagnetic media in addition to the above expressions. These equations are found to represent experimental facts fairly well except as regards magnetization in small fields. Without further knowledge of the nature, range, orientation of the internal field, however, it is not clear how the above is to be elaborated to give better agreement with experiment.

Another derivation is due to Heisenberg. From the Einstein-de Haas and Barnett Effects we know that ferromagnetism is essentially due to electron spins. Heisenberg showed that the same resonance forces that govern the coupling of spins in isolated atoms and molecules, and which had been studied in connection with the interpretation of spectra, were of a suitable order of magnitude to account for the internal fields of Weiss. These forces are known to be short range forces, so that to a first approximation it seemed desirable to assume that in ferromagnetic crystals only the spins of nearest neighbours interact. Ferromagnetism would then be associated with atoms whose configuration in the solid state was such that neighbouring antiparallel spins had a higher energy than neighbouring parallel spins. The energy of any given state of a crystal may then be written as a function of this number  $s$  of neighbouring antiparallel spins plus, if external magnetic fields are taken into account, a term involving the resultant magnetization. In order to subject this model to the appropriate statistical treatment and so calculate its most probable state under any given external conditions, it is necessary to know

the multiplicity of the state associated with each value of  $s$  and  $I$ . It has up to now been impossible to give the correct expression for the above-mentioned multiplicity, and in order to get around this difficulty Heisenberg made certain arbitrary assumptions which did in the end yield results essentially similar to equation (1) above. His work showed that resonance forces had in them stuff out of which ferromagnetism might be built, but it is probably not possible to conclude with any degree of certainty that the spin interactions give rise to a molecular field that appears in the equations in the simple way that Weiss assumed.

It seems fair to say, then, that equation (1) is not really a derived equation at all, but rather a semi-empirical equation, and further discussion is required to establish its theoretical significance.

We shall consider in some detail the following model.  $N$  magnetic elements, or briefly spins, are arranged in a geometrical array, and of these,  $n_+$  are pointing to the right, and  $n_-$  to the left. Only these two orientations are possible. Each spin interacts with its nearest neighbours only, of which we shall assume that there are  $z$ , and in such a way that if two neighbours are parallel their mutual energy is zero, if antiparallel, their mutual energy is  $2\epsilon/z$ . Then if  $y$  represents the magnetization measured on a scale for which absolute saturation is unity, we have the following:

$$n_+ + n_- = N \quad y = \frac{n_+ - n_-}{N} = \frac{2n_+ - N}{N}$$

$$\frac{n_+}{N} = \frac{1}{2}(1 + y) \quad \frac{n_-}{N} = \frac{1}{2}(1 - y).$$

We shall further use the following abbreviations:

$$\alpha = \frac{\mu H}{kT} \quad \beta = \frac{\epsilon}{kT}.$$

We shall also make use of the concept of a local intensity of magnetization. Of the  $z$  neighbours of any one spin,  $v_+$  are pointing to the right and  $v_-$  to the left, so that if  $i$  represents the relative intensity of magnetization of these  $z$  neighbours, we have

$$i = \frac{v_+ - v_-}{z} \quad \frac{v_+}{z} = \frac{1}{2}(1 + i) \quad \frac{v_-}{z} = \frac{1}{2}(1 - i).$$

Finally, if in any configuration of the whole system there are in all  $s$  pairs of reversed neighbours, the total energy of the state is given by

$$E(y, s) = -(n_+ - n_-) \mu H + \frac{2s\epsilon}{z} \quad (3)$$

and if  $p(y, s)$  is the number of ways in which the spins can be arranged among each other subject to the condition that the resultant magnetization is  $y$  and the number of antiparallel neighbours is  $s$ , then we have for the partition function

$$Z = \sum_{n_+, n_-, s} p(y, s) e^{-\frac{E(y, s)}{kT}}, \quad (4)$$

the summation being taken over all possible values of  $n_+$ ,  $n_-$ , and  $s$ .

The function  $p(y, s)$  can be given for two special cases only, namely, the linear chain of spins, for which  $z = 2$ , and a multidimensional crystal in which every atom is equidistant from every other one, and for which consequently  $z = N - 1 \sim \infty$ .

The first case,  $z = 2$ , has been studied by Ising\* who was able to give a complete solution of the problem. The result is a magnetization curve roughly similar to that usually obtained for paramagnetics, but with an increased susceptibility depending on  $z$ . The following item only is of interest in this discussion. If  $w(y)$  represents the probability of finding a magnetization of relative intensity  $y$ , where

$$w(y) = \frac{\sum p(y, s) e^{-\frac{E(y, s)}{kT}}}{Z}.$$

Ising finds that for  $H = 0$ ,  $w(y)$  has a sharp maximum at  $y = 0$ , or in other words that the only possible state in the absence of external fields is one of zero magnetization. Finite fields are required to produce a finite magnetization. The properties of the linear chain differ radically from those observed in typical ferromagnetics such as iron and nickel, both as regards thermal and magnetic characteristics.

The case  $z = N - 1 \sim \infty$  gives quite different results. Every arrangement of the  $n_+$  positive and  $n_-$  negative spins among each other gives a configuration for which  $s = n_+ n_-$ , and we may consequently write for  $p(y, s)$  simply the total possible number of re-arrangements of the spins, so that the partition function becomes

$$Z = \sum_{n_+, n_-} \frac{N!}{n_+! n_-!} e^{(n_+ - n_-)\alpha - \frac{2n_+ n_- \beta}{N-1}}. \quad (4A)$$

From this we find† for the magnetization  $y$ , the average energy per spin  $E$ ,

\* 'Z. Physik,' vol. 31, p. 253 (1925).

† The maximum term in the summation may be found, after substituting Stirling's expression for the factorials, by differentiating the logarithm of the general term above with respect to  $n_+$  and equating this to zero (condition for a maximum), giving

$$-\log n_+ + \log n_- + 2\alpha + \frac{2(n_+ - n_-)}{N} \beta = 0$$

$$\log \frac{n_+}{n_-} = \log \frac{1+y}{1-y} = 2(\alpha + \beta y)$$

$$y = \tanh(\alpha + \beta y).$$



the additional specific heat per spin  $\sigma$ , the critical or Curie temperature  $\theta$ , the jump in the specific heat  $\Delta\sigma$ , and the initial susceptibility  $\left[\frac{\partial y}{\partial H}\right]_{H=0}$ , the following expressions which are, moreover, exactly the formulæ of the usual Weiss Theory :

$$y = \tanh (\alpha + \beta y) \quad (5)$$

$$E = \frac{1}{2} (1 - y^2) \varepsilon, \quad (6)$$

$$\theta = \frac{\varepsilon}{k} = \beta T, \quad (7)$$

$$\sigma = k \frac{y^2 (1 - y^2) \beta^2}{1 - (1 - y^2) \beta}, \quad (8)$$

$$\Delta\sigma = \frac{3}{2} k, \quad (9)$$

$$\left[\frac{\partial y}{\partial H}\right]_{H=0} = \frac{\varepsilon}{\mu} \left(\frac{T}{\theta} - 1\right) \text{ for } T > \theta. \quad (10)$$

The partition function (4A) may be written

$$Z = \sum_{n_{\pm}} w(y),$$

where  $w(y)$  is as before the probability of finding the magnetization  $y$ . For temperatures below the Curie point and in the absence of external fields this function has two equal maxima. The sample may be spontaneously magnetized to the right or to the left, but it cannot be demagnetized. If the further assumption is made that the sample can somehow spontaneously jump from a positive to a negative magnetization, the above calculation of  $y$  from the partition function is no longer valid. We must rather evaluate  $Z$  and find  $y$  from the relation

$$y = \frac{\partial}{\partial x} \log Z,$$

which gives for small fields

$$y = \tanh \frac{I_w V H}{k T}.$$

( $I_w$  being the spontaneous magnetization given by  $\frac{I_w}{I_0} = \tanh \frac{I_w}{I_0} \beta$ ) or for infinite volumes  $V$  an infinitely steep magnetization curve. This, as we shall see further on, is probably the type of result which a correct theory would

give. Its derivation here is, however, not satisfactory because it comes from the assumption that only states for which  $y = \pm I_w/I_0$  are possible, and that spontaneous jumps from one to the other are responsible for the observed demagnetization. It is, however, distinctly encouraging to find that the outstanding flaw just mentioned is a consequence of precisely the assumption of  $N - 1$  neighbours which we shall have to modify. Consider the multi-dimensional crystal magnetized to the right, but subjected to a gradually increasing demagnetizing field. As the number of negative spins increases, the energy of the crystal increases according to equation (6). Further the very essence of this model is that its energy is independent of the precise arrangement of the reversed spins, so that there is no way out of the large energy required to reverse the direction of magnetization short of a simultaneous reversal of all  $N$  spins. In a three-dimensional crystal, however, the reversed spins can cluster together, and so reduce the energy of states of low magnetization, as the number of antiparallel neighbours in such a cluster will be restricted essentially to the spins at its boundary.

The essence of the problem, so far unsolved, is to find the most probable distribution of reversed spins in a three-dimensional crystal for any given external conditions. If the energy of the system were assumed independent of the arrangement of the reversed spins, we should have something very similar to an ideal gas. In fact, as will be shown below, the assumption of a random distribution leads to equations very similar to those found above for the multidimensional crystal. In reality, however, a random distribution cannot be the most probable one, at least at low temperatures, because the reversed spins attract each other. This attraction is merely another way of expressing the fact that the mutual energy between two reversed spins is less when they are neighbours than when they are separated. The reversed spins, therefore, must be treated as an imperfect gas, and it is to be expected that if their density is increased beyond a critical point at low temperatures, clustering, or as we should say for a gas, condensation and the formation of a new phase, is to be expected. The solution of our problem therefore involves the derivation of an equation of state from a knowledge of the interaction of the particles of the system. The difficulties of this problem are well known, and a complete solution is not to be expected.

The form of the solution, however, may be estimated by analogy with the problem of imperfect gases. The state of an imperfect gas is uniquely defined by two independent variables, for instance pressure and temperature, except for certain places on the phase diagram at which two or more phases can co-

exist. For instance, at the temperature  $T$  the whole sample is definitely in the vapour state if  $p < p_c$ , and definitely in the condensed state if  $p > p_c$ ,  $p_c$  being the vapour pressure at the temperature  $T$ . But for  $p = p_c$  any fraction of the sample may be in equilibrium in the vapour state. Similarly in our model of ferromagnetic crystals the reversed spins attract each other and may be considered capable of existing in two phases. For  $H > 0$  the positive spins are definitely in the condensed phase with the negative or reversed spins distributed more or less at random and forming a gaseous phase, etc. But for  $H = 0$  any portion of the negative spins may be in the condensed phase and any portion in the gaseous or dispersed phase. Consequently for  $H = 0$  the function  $w(y)$  should have not one, or two, maxima, as for  $z = 2$  and  $\infty$ , but should have an essentially flat top, or a maximum for any value of  $y$  between the limits  $-\frac{I_w}{I_0} \leq y \leq \frac{I_w}{I_0}$ . Only for small values of  $N$ , that is in small regions within the sample, should  $w(y)$  go over into the form with two maxima, corresponding to the existence of small regions of spontaneous magnetization.\* For  $H > 0$ , however, we may expect the reverse spins to be in the disperse phase, and may consequently treat them in a first approximation as an ideal gas.

It has not been proved, of course, that the answer to our problem will be of the above form, but this seems very likely, first, because we know that such phase changes are produced by short range attractive forces, as in gases; second, because this form of solution seems to lead to a satisfactory interpretation of experimental facts; and finally, because it explains why a rigorous statistical treatment of the three-dimensional model is so difficult, i.e., for the same reasons that it has so far been impossible to make a complete theory of imperfect gases.

The truth of the above statement that the Weiss equations follow from the assumption that the spins in a three-dimensional model of the type under consideration are arranged at random† may be seen from the following argument. Consider any one spin together with its  $z$  neighbours. The

\* The above analogy cannot be extended to include the latent heat of vaporization because in the ferromagnetic case the whole sample is in energetically the same state for the two phases corresponding to  $\pm I_w$ .

† More exactly, we shall assume that the probability that any one of the  $z$  neighbours of an atom be positive or negative is independent of the orientation of the other  $z - 1$  neighbours, which assumption may be expected to hold true approximately near saturation, especially for body centred cubic lattices such as iron ( $z = 8$ ) in which the nearest neighbours of an atom are not nearest neighbours to each other.

magnetization of the  $z$  neighbours is given by  $i$  as defined above, so that the two possible states of such a spin have the energies

$$-\mu H + \frac{2v_- \epsilon}{z}$$

or

$$\mu H + \frac{2v_+ \epsilon}{z}$$

From this follow immediately the expressions  $y(i)$  and  $E(i)$  for the most probable values of the magnetization and energy of the spin in question, the local magnetization being  $i$ .

$$y(i) = \frac{n_+ - n_-}{N} = \frac{ce^{\alpha - 2\frac{v_-}{z}\beta} - ce^{-\alpha - 2\frac{v_+}{z}\beta}}{ce^{\alpha - 2\frac{v_-}{z}\beta} + ce^{-\alpha - 2\frac{v_+}{z}\beta}} = \tanh(\alpha + \beta i), \quad (11)$$

and similarly

$$E(i) = \epsilon(1 - i \tanh i\beta). \quad (12)$$

The probability that any one of the  $z$  neighbours be a positive spin is  $n_+/N$ , and since by assumption the distribution of the positive and negative spins is a random one, the probability that any two spins be positive is simply  $(n_+/N)^2$ , and consequently the probability that of the  $z$  neighbours  $v_+$  are positive and  $v_-$  negative is

$$\left(\frac{n_+}{N}\right)^{v_+} \left(\frac{n_-}{N}\right)^{v_-}.$$

Finally, the arrangement of the positive and negative neighbours may be chosen in

$$\frac{z!}{v_+! v_-!}$$

ways, so that we may write for the probability of finding a local intensity  $i$  when the total magnetization is  $y$

$$\phi(i, y) = \frac{z!}{v_+! v_-!} \left(\frac{n_+}{N}\right)^{v_+} \left(\frac{n_-}{N}\right)^{v_-}.$$

The resultant magnetization and energy per spin of the total sample consequently are

$$y = \sum_i \phi(i, y) y(i) \quad (13)$$

$$E = \frac{1}{z} \sum_i \phi(i, y) E(i), \quad (14)$$

the summation extending over all possible values of  $i$ . The factor  $\frac{1}{z}$  must be introduced into the expression for the energy as we have counted the mutual

energy of each oppositely directed pair of spins twice, once with each member of the pair. For very large values of  $z$  these expressions reduce to the corresponding expressions (5) and (6) above.

Equations (13) and (14) are very cumbersome for numerical calculating, and do not give results radically different from the simpler expressions of the Weiss theory. By way of illustration, the case  $z = 8$  has been computed. This corresponds to a body centred cubic lattice, such as  $\alpha$ -iron has. The results are plotted in fig. 1 showing the magnetization in zero field as given by

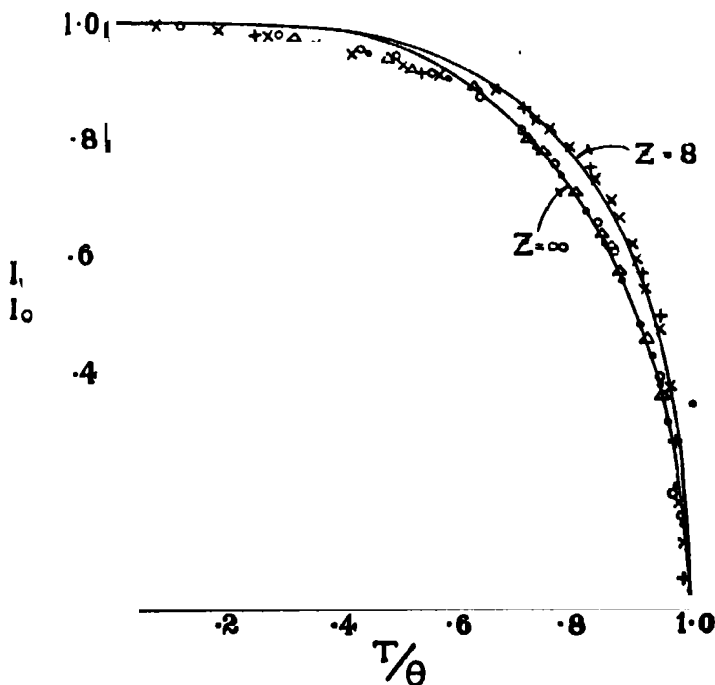


FIG. 1. ● Nickel, Weiss and Forrer; ○ nickel, Bloch; Δ cobalt, Bloch; × iron, Hegg; + iron, Curie.

equation (13), and the corresponding result according to equation (5) for  $z = \infty$ . The computations for  $z = 12$  (Ni and Co) were not carried out, but the curve for this would be somewhere between the two shown, roughly half-way between. The experimental observations had been collected for a previously published article\* to which the reader is referred for further comment. The experimental points just below the Curie point do show a larger spontaneous magnetization for  $z = 8$  than for  $z = 12$  in agreement with the theory. Exact agreement throughout cannot be expected because the theory assumes one

\* Bitter, 'Phys. Rev.,' vol. 39, p. 340 (1932).

magnetic element per atom, a condition not met with in either Ni, Co, or Fe, for which one obtains approximately  $3/5$ ,  $3/2$ , and 2 Bohr magnetons per atom respectively from measurements on absolute saturation; and secondly, because there is no *a priori* justification for the assumption that the reversed spins may be treated as an absolutely perfect gas even in magnetically saturated states. In fact, the (usually slight) discrepancy between theory and experiment as regards saturation at very low temperature, the specific heat, and the initial susceptibility above the Curie point, all indicate that the assumptions are not completely fulfilled. We may conclude, however, that the qualitative agreement throughout indicates that to a first approximation the reversed spins may be treated as a perfect gas, provided that their density is less than that corresponding to the saturation intensity for any temperature below the Curie point, and for any density at all above the Curie point. Or we might say that fig. 1 may be interpreted as a phase diagram, the magnetization being equivalent to the usual density, and the area under the curve representing states for which the condensed phase for reversed spins exists. Hysteresis effects then follow as being due to obstacles obstructing the growth of the new phase,\* or to other causes tending to produce supersaturation.

According to the above the magnetization of a perfect single crystal should proceed to saturation in the direction of easy magnetization in infinitesimal fields, the value of the saturation intensity being given by the "vapour pressure" of the spin gas. Actual crystals do not behave in this way. Experimental results for the magnetization of iron crystals along a [100] axis are shown in fig. 2. It is not sufficient to say that hysteresis effects are excluded in the above statement of the problem, and that as hysteresis effects are prominent in small fields, no agreement is to be expected. On the curve in question the hysteresis loop is approximately as wide as the size of the points on the diagram indicate, so that we must account for considerable reversible departures from the ideal curve for perfect crystals. Actual crystals are certainly not geometrically perfect. Because of impurities, because of strains produced in handling, machining, etc., or in the absence of such accidentally produced distortions, because of the thermal vibration of the lattice, small regions in some parts of the crystal will have a slightly different symmetry from small regions in other parts. The effect of such lattice distortions can be calculated by means of Becker's Theory. The exact calculations are very tedious, involving a great deal of graphical integration, and only a simple approximation will be attempted below.

\* This point has been discussed by Bloch, 'Z. Physik,' vol. 74, p. 295 (1932).

In a perfect single crystal the direction of magnetization is given according to Akulov by the values of  $\alpha_i, \alpha_j, \alpha_k$ , for which the expression

$$c \sum_{i,j,k} \alpha_i^2 \alpha_j^2 - \vec{I}_w \cdot \vec{H}$$

is a minimum,  $\alpha_i$ , etc., being the direction cosines of the magnetization  $\vec{I}_w$  with respect to rectangular co-ordinates parallel to the cubic axes of the crystal, and  $c$  a constant which is

$$\text{for Fe} \quad c = 2 \cdot 15 \times 10^5,$$

$$\text{for Ni} \quad c = -2 \cdot 8 \times 10^4.$$

For homogeneously distorted crystals the expression to be minimized is

$$E_0 = c \sum \alpha_i^2 \alpha_j^2 + K_1 \sum A_{ii} \alpha_i^2 + K_2 \sum A_{ij} \alpha_i \alpha_j - \vec{I}_w \cdot \vec{H}, \quad (15)$$

where  $A_{ij}$  is the tensor describing the lattice distortion, and  $K_1$  and  $K_2$  are constants derivable from observations on magnetostriction.\*

In a randomly distorted crystal the tensor  $A_{ij}$  will be a function of the space co-ordinates, and the magnetization curve of the whole crystal will be the average of the magnetization curves for all the small distorted regions. For small fields and small distortions, the minima of  $E_0$  are given essentially by the first term in equation (15), so that for the initial part of the magnetization curve the only values of  $\alpha_i$ , etc., to be considered are  $\pm 1$  or 0 for iron and  $\pm 1/\sqrt{3}$  for nickel. In iron, for instance,  $E_0$  for an undistorted single crystal in the absence of external fields has equal minima along the  $i, j$ , and  $k$ -axes. The effect of lattice distortions will be primarily to make in some parts of the crystal the minima along the  $i$ -axis deeper than those along the other axes, and similarly for the minima along the  $j$  and  $k$ -axes in other parts of the crystal. Therefore, provided only that the distortions are distributed at random, their effect will be to limit the magnetization parallel to a  $[100]$  axis in infinitesimal fields to  $\frac{1}{3}\vec{I}_w$ , because over  $2/3$  of the crystal magnetization along the  $\pm j$  and  $\pm k$ -axes will be energetically preferable, and infinitesimal fields will therefore be unable to change the direction of magnetization from  $\pm j$ , or  $\pm k$  to  $+i$ . We may thus expect a first break in the magnetization curve for  $I_1 = \frac{1}{3}\vec{I}_w$ ,  $H_1 = 0$ . Similarly if an infinitesimal magnetic field is applied along the diagonal axis in the  $j-k$  plane midway between the  $+j$  and  $+k$ -axes, one-third of the crystal will be magnetized along the  $+j$  and one-third along the  $+k$ -axes, the remaining third being magnetized along the

\* Bitter, 'Metallwirtschaft,' vol. 12, pp. 720, 735 (1933).

$\pm i$ -axis which is perpendicular to  $H$ . The resultant component of  $I_w$  in the direction of  $H$  for such a configuration is evidently  $I_1 = \frac{2}{3} 1/\sqrt{2} I_w$  or  $0.472 I_w$ . And similarly for a field parallel to a trigonal axis we have  $I_1 = 1/\sqrt{3} I_w$ . These results, together with others arrived at by similar reasoning for nickel, are contained in Table I.

Table I.

Direction of magnetization.	Iron.				Nickel.			
	$H_1$ .	$I_1/I_w$ .	$c'H_1$ .	$I_1/I_w$ .	$H_1$ .	$I_1/I_w$ .	$c'H_1$ .	$I_1/I_w$ .
[100]	0	1/3	1	1	0	$1/\sqrt{3}$	0	$1/\sqrt{3}$
	0	0.333	1	1	0	0.58	0	0.58
[110]	0	$\sqrt{2}/3$	$\sqrt{2}$	$1/\sqrt{2}$	0	$\frac{1}{2}\sqrt{\frac{2}{3}}$	$\sqrt{\frac{2}{3}}$	$\sqrt{\frac{2}{3}}$
	0	0.472	1.41	0.71	0	0.408	1.23	0.816
[111]	0	$1/\sqrt{3}$	0	$1/\sqrt{3}$	0	$\frac{1}{2} + \frac{1}{2} \cos \alpha$	$1/1 - \cos \alpha$	1
	0	0.58	0	0.58	0	0.498	1.5	1

$$\alpha = \text{angle between trigonal axes} = 2 \cos^{-1} \sqrt{\frac{2}{3}}.$$

A second break in the magnetization curve is to be expected when the difference in magnetic energy along two axes becomes greater than the energy difference due to strains. Assuming that the strains may produce a difference in energy for magnetization along two axes never exceeding  $\epsilon$ , it follows that saturation, in iron along a [100] axis for instance, will be reached when  $I_w H_2 = \epsilon$ , or  $c'H_2 = 1$ , where  $c' = I_w/\epsilon$ .

Similarly, when a field of intensity  $H_2 = \sqrt{2}/c'$  is applied along a diagonal axis in the  $j-k$  plane of an iron crystal, the whole of the crystal will be magnetized along the nearest directions of easy magnetization, the resultant intensity at this point being  $I_2 = I_w/\sqrt{2}$ . The results of similar calculations for the second breaks in the various magnetization curves are contained in Table I. Briefly, the modifications in the ideal magnetization curves of single crystals introduced by the assumption of lattice imperfections are the following.

In perfect crystals the magnetization of the entire crystal goes to saturation in infinitesimal fields in the direction of easy magnetization nearest the applied field.

In imperfect crystals having  $n$  directions of each magnetization, a fraction  $1/n$  of the total crystal will be magnetized to saturation in each of the  $n$  directions



in the energetically most favourable sense in infinitesimal fields. A finite field, whose strength is dependent on the lattice distortion and the orientation of the crystal, is required to magnetize the entire crystal in that particular direction of easy magnetization nearest to the applied field.

The ideal magnetization curve of imperfect crystals will therefore proceed in infinitesimal fields to the first critical point at  $I_1$ , then at a different rate to

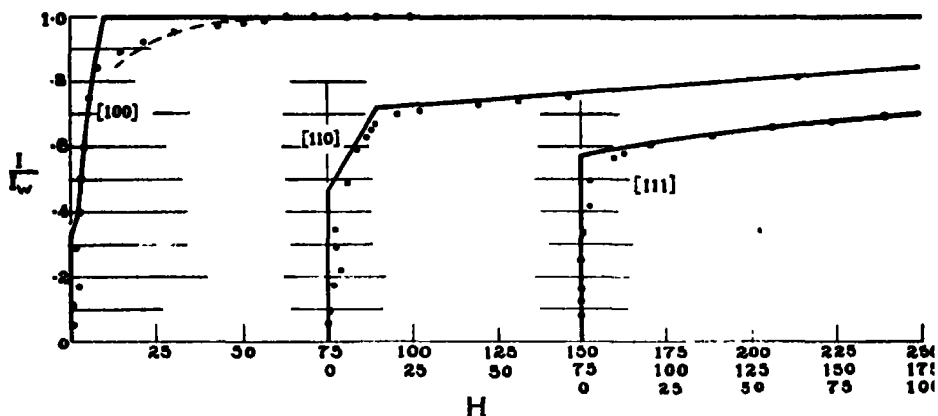


FIG. 2.—Magnetization curve for iron.

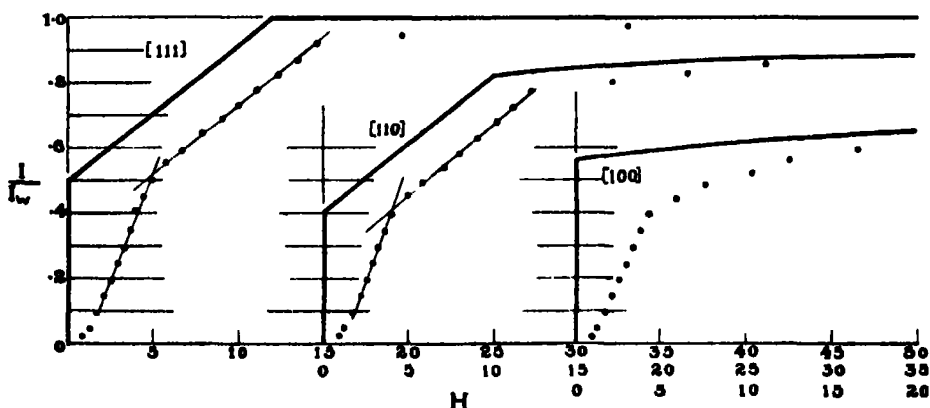


FIG. 3.—Magnetization curve for nickel.

the second critical point at  $I_2$ ,  $H_2$ , and finally will approach saturation due to the rotation of the directions of easy magnetization in large fields as in Akulov's Theory. Such curves are plotted for iron and nickel in figs. 2 and 3 respectively, the only disposable constant  $c'$  being chosen  $1/10$  for iron and  $1/12$  for nickel. The two critical points have been arbitrarily connected by a straight

line except for iron magnetized parallel to a tetragonal axis. In this case an actual calculation of the curve was carried out assuming  $c = \infty$  and assuming that the terms in  $E_0$  involving the distortion represented an oblate ellipsoid randomly oriented in the crystal. Representative magnetization curves for various orientations of the ellipsoid were plotted, and their average found by graphical methods. The experimental points in figs. 2 and 3 were taken from papers by Honda and Kaya\* and Kaya.† The experimental results for iron are not very accurate, and the best that one can say is that they are not in disagreement with the theoretical curves except as regards the corners at  $I_2$ ,  $H_2$ . This will be discussed further on. It should be remarked that steeper curves than those plotted have been observed both by Honda, Masumoto, and Kaya,‡ and by Dussler and Gerlach,§, which indicates that the strains responsible for the observations of fig. 2 are accidentally produced in the preparation of the sample.

The experimental results for nickel are much more accurate and could therefore be plotted on a larger scale, which, in turn, enlarges the hysteresis effects. The agreement between theory and experiment, such as it is, leaves little doubt that the suggested mechanisms do actually play a part in the magnetization of nickel crystals.

A similar comparison between theory and experiment should be made for the magnetostriction of single crystals as a function of the applied field, but the data at present available are too inaccurate to afford a reasonable check.

The sharp corners in the magnetization curves drawn in figs. 2 and 3 at  $I_2$ ,  $H_2$  are a consequence of the assumption that  $\alpha_i$ , etc., can have only discrete values, or in other words that  $c$  in equation (15) may be treated as infinite. The fact that in actual crystals  $c$  is finite means that owing to the crystal imperfections the magnetization will not be confined to certain crystallographic directions even in infinitesimal fields, but that there will be a certain probability of finding regions within the crystal for which the directions of easy magnetization differ from the corresponding directions in perfect crystals. The amount of rounding off of the corners which this produces follows without further assumption from an exact computation of the curves in question, but these computations are, as has been said, rather involved, and only the simple case  $c = 0$  has been worked out. The results obtained are probably a fair

\* 'Sci. Rep. Tohoku Imp. Univ.,' vol. 15, p. 721 (1926).

† 'Sci. Rep. Tohoku Imp. Univ.,' vol. 17, p. 639 (1928).

‡ 'Sci. Rep. Tohoku Imp. Univ.,' vol. 17, p. 111 (1928).

§ 'Z. Physik,' vol. 44, p. 284 (1927).

approximation to the approach to saturation, for instance in iron in a [100] direction.

For  $c = 0$  equation (15) may be written, assuming the magnetic field parallel to the  $k$ -axis,

$$E_\theta = \epsilon \sin^2 (\psi - \theta) - I_w H \cos \theta, \quad (16)$$

$\theta$  being the angle between the  $k$ -axis and the magnetization, and  $\psi$  the angle between the direction of easy magnetization produced by the strains and the  $k$ -axis, it being assumed that the strains produce one direction of each magnetization, and that the energy required to rotate the magnetization through  $90^\circ$  from the direction of easy magnetization is  $\epsilon$ . The magnetization curve of a region having a distortion such as to produce the conditions described in equation (16) is

$$H = \frac{2\epsilon}{I_w} \frac{\sin (\psi - \theta) \cos (\psi - \theta)}{\sin \theta}$$

or solving for  $I_w \cos \theta = I$

$$I = f(H, \epsilon, \psi).$$

The probability that strains randomly orientated will produce a direction of easy magnetization between  $\psi$  and  $\psi + d\psi$  is  $\sin \psi d\psi$ , so that we have for the magnetization of a randomly distorted crystal

$$I = \int_0^\pi f(H, \epsilon, \psi) \sin \psi d\psi. \quad (17)$$

The curve representing equation (17) was calculated graphically and is shown as a broken line in fig. 2, the same constant  $\epsilon = I_w/c' = 10I_w$  being used as for the other curves for iron. The agreement is sufficiently good to indicate that randomly oriented imperfections are concerned in producing the observed approach to saturation.

A rough estimate of the order of magnitude of the strains necessary to account for the observed values of  $c'$  in figs. 2 and 3 indicates that for nickel these are probably slightly larger than the magnetostrictive strains, being from 5 to  $10 \times 10^{-6}$ , and for iron from 5 to 10 times larger.

In ferromagnetic materials having no magnetostriction the constants  $K_1$  and  $K_2$  in equation (15) vanish, so that the strains should have no effect on the magnetization curve. Observations on such materials should approach more closely the square curves of Akulov's Theory. In permalloy, an iron-nickel alloy containing roughly 20% of iron, the constant  $c$  as well as  $K_1$  and  $K_2$  approach zero, so that even a polycrystalline wire should show the characteristic behaviour of a perfect single crystal in the direction of each magnetization,

as it in fact does. A curve published by Buckley and McKeehan\* shows practically complete saturation in fields of 5 or 6 oersteds, whereas in iron and nickel crystals fields of about 10 times this intensity are required.

It is a great pleasure to acknowledge my indebtedness to many of the physicists at the Cavendish Laboratory, especially Professor R. H. Fowler, for the benefit of stimulating discussions on the subject treated in this paper, and to the Guggenheim Foundation and the Westinghouse Electric and Manufacturing Company for making my stay in Cambridge possible.

### *Summary.*

The statistical theory of spontaneous magnetization is discussed based on a model consisting of a geometrical array of magnetic elements, of which only the nearest neighbours interact. It is shown that the usual Weiss-Heisenberg equation is the rigorous solution of this model, assuming that each element is the nearest neighbour to every other element. The correct solution for actual three-dimensional crystals cannot be given, but very probably is characterized by an infinitely steep magnetization curve in small fields up to a saturation value given by the Weiss equation. A new treatment applicable to the saturation part of the magnetization curve gives approximately the observed dependence on the number of nearest neighbours at not too low temperatures. Finally an attempt is made to show quantitatively that the discrepancies between the magnetization curves predicted above and the curves observed for actual iron and nickel crystals in small fields are due to crystal imperfections.

\* 'Phys. Rev.,' vol. 26, p. 261 (1925).

---

## *A Relativistic Basis of the Quantum Theory.—II.*

By H. T. FLINT, Reader in Physics, University of London, King's College.

(Communicated by O. W. Richardson, F.R.S.—Received February 15, 1934.)

This paper is a continuation of the last paper communicated to these 'Proceedings.' In that paper, which we shall refer to as the first paper, a more general expression for space curvature was obtained than that which occurs in Riemannian geometry, by a modification of the Riemannian covariant derivative and by the use of a fifth co-ordinate.

By means of a particular substitution

$$\left( \Lambda_{\mu}{}^{\sigma} = \frac{1}{\psi} \frac{\partial \psi}{\partial x^{\mu}} \right)$$

it was shown that this curvature takes the form of the second order equation of quantum mechanics. It is not a matrix equation, however, but one which has the character of the wave equation as it occurred in the earlier form of the quantum theory. But it contains additional terms, all of which can be readily accounted for in physics, except one which suggested an identification with the energy of the spin.

Towards the end of the paper the relation of the matrices to the theory was discussed and we shall now consider the matrix theory in the light of the suggestions made.

The chief points of the present contribution to the subject are to show how Dirac's equations fit into the theory and to derive the second order equation in its matrix form. This is not derived by a double operation of the first order operator, but merely by a substitution of a relation, which these equations provide. The method of the former contribution is extended to matrices and the substitution referred to above is replaced by its counterpart in the matrix theory.

### *The First Order Equations.*

It will be convenient to write down the equations of Dirac for comparison.

For the electron in the absence of an electromagnetic field they are written in the matrix form

$$\left( \frac{W}{c} + \alpha_x p_x + \alpha_y p_y + \alpha_z p_z + \alpha_m m c \right) \psi = 0, \quad (1)$$

where

$W, p_x, p_y, p_z$  are the operators:  $\frac{h}{2\pi} \frac{\partial}{\partial t}, \frac{h}{2\pi i} \frac{\partial}{\partial x},$  etc.,

$h$  being Planck's constant.

We shall write this in the form

$$\alpha_\mu p_\mu \psi = 0, \quad (2)$$

where  $\mu$  has the values 1 — 5, and  $\alpha_5$  takes the place of  $\alpha_m$ ,  $p_\mu$  of  $mc$ .

This notation places all the quantities on the same footing and makes  $mc$  a momentum, apparently without a co-ordinate as its conjugate.

This raises again the difficulties associated with these equations. In Dirac's theory all these quantities,  $p$ , are not treated in the same way,  $mc$  is treated as a number and the time  $t$  is not treated exactly as the other co-ordinates. These points have been discussed in some detail in a recent paper by Proca.\*

We shall here regard these quantities as equivalent in the relativistic sense and associated with conjugate co-ordinates  $x_1, \dots, x_5$ . The last of these co-ordinates has not yet been interpreted in physics and it has hitherto been necessary to remove it from any equations which have to be applied to practical problems. Its use has been required in a relativistic theory of Kaluza and Klein and it seems to be required here as a conjugate to  $p_5$ .

We shall develop a general theory and, as mentioned in the last paper, we shall suppose that the existing quantum theory is, as it were, an approximation which can be attained by giving special privileges to  $p_5$  and  $x_5$ .

The extension to include an electron in an electromagnetic field is made by means of the equation

$$\left\{ \frac{W}{c} + \frac{e}{c} \phi_0 + \alpha_m \left( p_m + \frac{e}{c} \phi_m \right) + \alpha_5 mc \right\} \psi = 0, \quad (3)$$

which we shall write

$$\alpha_\mu \left( p_\mu + \frac{e}{c} \phi_\mu \right) \psi = 0, \quad (4)$$

where  $(p_5 + [e/c] \phi_5)$  replaces  $mc$  and the other quantities,  $p$ , are regarded as operators.

The quantities,  $\phi$ , are the potentials of the electromagnetic theory and we shall write  $\phi_5 = 0$ .  $p_5$  is an operator and the passage from (4) to (3) implies that  $p_5 \psi$  or  $\frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial x^5} = mc$ . This result is made use of in the theory of Klein, though not in a form proposed by de Broglie. The latter writes

$$\frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial x^5} = \frac{e}{\alpha} \psi.$$

\* 'Ann. Physique,' vol. 20, p. 348 (1933).

The assumption  $\alpha = e/mc$  has certain interesting consequences of interest in the quantum theory which have been pointed out before.\*

In papers referred to in the first paper it has been pointed out that the first order equations can be written in the form of a vanishing divergence. This view suggests that here they should be

$$\frac{\partial}{\partial x^\mu} (\alpha^\mu \psi) + \Gamma_{\beta\mu}{}^\alpha \alpha^\beta \psi = 0, \quad (5)$$

where the  $\Gamma_{\beta\mu}{}^\alpha$  is the Christoffel bracket expression of Riemannian geometry for a five-dimensional continuum.

But  $\alpha^\mu$  must satisfy the variation

$$\frac{\partial \alpha^\mu}{\partial x^\mu} + \Gamma_{\beta\mu}{}^\alpha \alpha^\beta + T_{\beta\mu}{}^\alpha \alpha^\beta = 0 \quad (6)$$

(see first paper).

This converts (5) into

$$\alpha^\mu \frac{\partial \psi}{\partial x^\mu} - T_{\beta\mu}{}^\alpha \alpha^\beta \psi = 0. \quad (7)$$

According to the first paper

$$T_{\mu\nu}{}^\sigma = \Lambda_{\mu\nu}{}^\sigma - \gamma^{\sigma\tau} (\gamma_{\mu\rho} \Lambda_{\tau\nu}{}^\rho + \gamma_{\nu\rho} \Lambda_{\tau\mu}{}^\rho), \quad (8)$$

consequently

$$T_{\beta\mu}{}^\alpha = 2\Lambda_{\beta\mu}{}^\alpha. \quad (9)$$

In former papers† reasons have been given for an identification of  $2\Lambda_{\beta\mu}{}^\alpha$  with  $\phi_\beta$  or possibly  $(p_\beta + [e/c] \phi_\beta)$ .

Comparison of (7) with (4) shows at once that it may be identified with Dirac's equation for an electron in an electromagnetic field. Equation (7) can be interpreted as an equation in a function  $\psi$  which causes the matrices to behave like contravariant quantities in Riemannian geometry. This gives what we may call an analytical meaning of  $\psi$ .

It is true that this contravariant property is given to the  $\alpha$ 's in a limited sense. According to (5) it makes the vanishing of the divergence Riemannian in type. In former discussions on this subject a departure had to be made from this relatively simple type of differentiation.

There is no doubt about the importance of this type of variation in physics. We may say that Nature seems to make use of it whenever possible and, even when compelled to modify it in the domain where quantum mechanics prevails, she introduces  $\psi$  in order to get back to it.

\* Flint, 'Proc. Roy. Soc.,' A, vol. 126, p. 644 (1930).

† Flint, 'Proc. Roy. Soc.,' A, vol. 121, p. 676 (1928); vol. 141, p. 363 (1933).

There has always been a difficulty in discovering exactly what place these equations occupy in any relativistic theory. Equation (7), which is the version of them in the present theory, is an equation for  $\psi$ , which gives the vanishing divergence (5), and (7) is the matrix counterpart of the substitution of the first paper. Instead of

$$\partial\psi/\partial x^\mu = \Lambda_{\mu\sigma}{}^\sigma \psi$$

we substitute

$$\alpha^\mu \frac{\partial\psi}{\partial x^\mu} = T_{\mu\sigma}{}^\sigma \alpha^\mu \psi. \quad (7)$$

*The Interpretation of  $T_{\mu\sigma}{}^\sigma$ .*

If (7) be compared with (4), we obtain the value of  $T_{\mu\sigma}{}^\sigma$ , which makes it equivalent to the quantum equations.

Remembering that in these equations  $p_\mu = \frac{\hbar}{2\pi i} \frac{\partial}{\partial x^\mu}$ , it follows that

$$T_{\mu\sigma}{}^\sigma = 2\Lambda_{\mu\sigma}{}^\sigma = -\frac{2\pi i}{\hbar c} e\phi_\mu.$$

To make (7) equivalent to Dirac's equations we must identify the first four components  $\phi_1, \dots, \phi_4$  ( $\phi_m$  typically) with the electromagnetic potential. We must give the value zero to  $\phi_5$  and replace  $\partial/\partial x^5$  by  $2\pi imc/\hbar$ .

This, of course, gives a special privilege to  $x^5$ , which is not enjoyed by the other co-ordinates, the operator  $\partial/\partial x^5$  being replaced by a number.

A similar suggestion with regard to  $\Lambda_{\mu\nu}{}^\nu$  was made in an earlier paper\* in a derivation of the second order equation for the electron without spin. There the number of co-ordinates was four.

In extending the discussion to five co-ordinates the quantity,

$$\Pi_\mu = p_\mu + (e/c) \phi_\mu,$$

replaces  $p_\mu$ , the momentum, which is absorbed into the quantity. The suggestion arises that  $\Pi_\mu$  is to replace  $\phi_\mu$  in these relations,† and one may readily incline to the belief that  $T_{\mu\sigma}{}^\sigma$  should be replaced by  $-(2\pi i/\hbar) \Pi_\mu$ .

If this be a correct identification, then (7) becomes

$$\alpha^\mu \frac{\partial\psi}{\partial x^\mu} + \frac{2\pi i}{\hbar} \Pi_\mu \alpha^\mu \psi = 0. \quad (7)$$

The question whether such an equation could be adopted in physics is not discussed here. Dirac's equations fit in so well with experiment that additional

\* Flint, 'Proc. Roy. Soc.,' A, vol. 121, p. 690 (1928).

† Flint, 'Proc. Roy. Soc.,' A, vol. 141, p. 373 (1933).



terms might be ruled out on experimental grounds. We see from this discussion that the equations of the quantum theory bear a close relation to an equation which has a simple meaning in a five-dimensional continuum and which seems satisfactory on relativistic grounds. The difficulties associated with the quantum equations arise as a consequence of our ignorance of the nature of  $x^5$ . It is, as stated above, as if we had to make an approximation to get rid of this co-ordinate. This and the way the term  $mc$  is introduced may mean that we are also uncertain about the nature of  $m$  (the rest mass) itself and that something is escaping us in this connection.

### *The Second Order Equations.*

The method of arriving at an equation of the second order is usually to make a double application of the operator which acts upon  $\psi$  in the equations (3) or (7). In this way a generalization of Laplace's operator is obtained. The process is a somewhat arbitrary one, and in some contributions to the 'Proceedings' it has been replaced by one which belongs naturally to the analysis employed.

In the first paper a second order equation was obtained from the expression for the curvature by the substitution mentioned. The form obtained was remarkable, for it contained terms which corresponded exactly to the requirements of the quantum theory. Identification of the terms presented no difficulty except in one case, the appearance of which offered no obvious clue, but the suggestion was made that it represented the spin-energy term. This is not quite an arbitrary suggestion for earlier work suggests that in an equation of the simple Schroedinger type, a term of this kind must occur.

The curvature equation obtained before making the substitution is

$$P_{\mu\nu} = R_{\mu\nu} - T_{\mu\nu}{}^{\sigma}{}_{;\sigma} + T_{\mu\sigma}{}^{\sigma}{}_{;\nu} - T_{\rho\sigma}{}^{\sigma} T_{\mu\nu}{}^{\rho} + T_{\rho\nu}{}^{\sigma} T_{\mu\sigma}{}^{\rho}, \quad (10)$$

where  $T_{\mu\nu}{}^{\sigma}{}_{;\sigma}$  denotes covariant differentiation in the sense of Riemannian geometry.

With the form adopted for  $T_{\mu\nu}{}^{\sigma}$  (8), it was shown that (10) gives

$$\begin{aligned} \gamma^{\mu\nu} P_{\mu\nu} - \gamma^{\mu\nu} R_{\mu\nu} = & 4\gamma^{\mu\nu} \Lambda_{\mu\sigma}{}^{\sigma}{}_{;\nu} + 4\gamma^{\rho\tau} \Lambda_{\rho\sigma}{}^{\sigma} \Lambda_{\tau\alpha}{}^{\alpha} \\ & + \gamma^{\rho\tau} (\gamma^{\sigma\beta} \gamma_{\sigma\alpha} + 2\delta_{\sigma}{}^{\sigma} \delta_{\alpha}{}^{\beta}) \Lambda_{\tau\sigma}{}^{\sigma} \Lambda_{\beta\rho}{}^{\rho}. \end{aligned} \quad (11)$$

The substitution  $\Lambda_{\mu\sigma}{}^{\sigma} = \frac{1}{\psi} \frac{\partial \psi}{\partial x^{\sigma}}$  gives the result we have described as the second order equation, but the introduction of the matrices and the relation of

(7) with its interpretation of the quantity  $\psi$ , suggests the substitution of this non-matrix relation by (7).

It is an advantage to get rid of this simpler substitution, for if  $\Lambda_{\mu\sigma}{}^\sigma$  is a constant factor multiplied by  $\phi_\mu$ , we cannot make  $\phi_\mu$  a component of a gradient, since this implies the absence of an electromagnetic field.

The difficulty is not so great if  $\phi_\mu$  is replaced by  $\Pi_\mu$  and this point has been discussed previously.\* We have to remember that part of  $\Pi_\mu$ , viz.,  $p_\mu$ , is of the form  $\partial W/\partial x^\mu$  in Mechanics, and that if  $\Pi_\mu$  is of the same form, so also is  $\phi_\mu$ . It would appear that this is not a possible general substitution for  $\Pi_\mu$  but that it denotes a selection of certain paths in the continuum, corresponding to the quantum integral  $\int p dq = n\hbar$ .

We can remove the difficulty here by making use of the matrix substitution.

We could substitute immediately the value proposed for  $\Lambda_{\mu\sigma}{}^\sigma$ ,  $\Lambda_{\mu\sigma}{}^\sigma \propto \phi_\mu$  or  $\propto \Pi_\mu$ , and obtain an equation for the quantity  $P = \gamma^{\mu\nu} P_{\mu\nu}$ . Such an equation may be of importance, but it is not a wave equation such as we are now seeking and since we know of no other relation for  $\psi$  than (7) we must use it in some way.

Moreover, we do not know a value of  $\Lambda_{rs}{}^\sigma$  except in terms of the matrices,  $\alpha_\mu$ . The connection between these quantities may be obtained from

$$\frac{\partial \alpha_\mu}{\partial x^\nu} = \Gamma_{\mu\nu}{}^\beta \alpha_\beta + T_{\mu\nu}{}^\beta \alpha_\beta \quad (12)$$

which, according to the first paper, is satisfied by the matrices. Evidently

$$\frac{\partial \alpha_\mu}{\partial x^\nu} - \frac{\partial \alpha_\nu}{\partial x^\mu} = (T_{\mu\nu}{}^\beta - T_{\nu\mu}{}^\beta) \alpha_\beta = 2\Lambda_{\mu\nu}{}^\beta \alpha_\beta. \quad (13)$$

This connection between the quantities occurring in (11) and the matrices, makes it natural to consider the expression:  $\gamma^{\mu\nu} P_{\mu\nu}$ , which occurred in the first paper and to substitute matrix expressions in the resulting expression.

It is not difficult to show from (8) that

$$\gamma^{\mu\nu} T_{\mu\nu}{}^\sigma = -\gamma^{\sigma\tau} T_{\tau\rho}{}^\rho, \quad (14)$$

so that by (10)

$$\gamma^{\mu\nu} (P_{\mu\nu} - R_{\mu\nu}) = 2 (\gamma^{\sigma\tau} T_{\tau\rho}{}^\rho)_{;\sigma} + \gamma^{\sigma\tau} T_{\tau\sigma}{}^\sigma T_{\rho\sigma}{}^\rho + \gamma^{\mu\nu} T_{\mu\nu}{}^\sigma T_{\mu\sigma}{}^\rho. \quad (15)$$

We can introduce matrices through the  $\gamma^{\mu\nu}$  in the following way: the product

\* Flint, 'Proc. Roy. Soc.,' A, vol. 177, p. 630 (1928).

$\alpha^\mu \alpha^\nu$  can be divided into a symmetrical and an anti-symmetrical part, which we can write in the form

$$\alpha^\mu \alpha^\nu = \gamma^{\mu\nu} + s^{\mu\nu}. \quad (16)$$

Now  $\gamma^{\mu\nu}$  is a particularly simple type of matrix and means simply the  $\gamma^{\mu\nu}$  of the line element, i.e., it is the contravariant of  $\gamma_{\mu\nu}$  in the expression  $\gamma_{\mu\nu} dx^\mu dx^\nu$ , multiplied by the unit matrix. If we assume the unit matrix to be present along with this component, there is no need to make any further distinction in the notation.

Thus the symmetrical part of  $\alpha^\mu \alpha^\nu$  is a diagonal matrix.

This is only another way of stating Tetrode's relations and the  $s^{\mu\nu}$  is the five-dimensional extension of the  $s^{\mu\nu}$  of Schroedinger's\* work.

We have similar covariant quantities

$$\alpha_\mu \alpha_\nu = \gamma_{\mu\nu} + s_{\mu\nu}. \quad (17)$$

If we take the term  $(\gamma^{\sigma\tau} T_{\tau\rho}{}^\rho)_\sigma$  from (15), we see that it may be written

$$\{(\alpha^\sigma \alpha^\tau - s^{\sigma\tau}) T_{\tau\rho}{}^\rho\}_\sigma.$$

If we make use of the relation corresponding to (12) viz.,

$$\frac{\partial \alpha^\mu}{\partial x^\sigma} = -\Gamma_{\beta\sigma}{}^\mu \alpha^\beta - T_{\beta\sigma}{}^\mu \alpha^\beta, \quad (18)$$

and of the following, which may be deduced from it

$$s^{\mu\nu}{}_{;\lambda} = -T_{\beta\lambda}{}^\mu s^{\beta\nu} - T_{\beta\lambda}{}^\nu s^{\mu\beta}, \quad (19)$$

we obtain instead of this term the expression

$$\alpha^\sigma (\alpha^\tau T_{\tau\rho}{}^\rho)_\sigma - \alpha^\beta \alpha^\tau T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho + s^{\beta\tau} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho + s^{\sigma\beta} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho - s^{\sigma\tau} T_{\tau\rho}{}^\rho{}_{,\sigma}. \quad (20)$$

On account of the anti-symmetric property of  $s^{\beta\tau}$ , the third term is zero, as is also part of the second. Thus (20) simplifies to

$$\alpha^\sigma (\alpha^\tau T_{\tau\rho}{}^\rho)_\sigma - \gamma^{\beta\tau} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho + s^{\sigma\beta} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho - s^{\sigma\tau} T_{\tau\rho}{}^\rho{}_{,\sigma}. \quad (21)$$

Thus all the terms on the right of (15) taken together give

$$2\alpha^\sigma (\alpha^\tau T_{\tau\rho}{}^\rho)_\sigma + 2s^{\sigma\beta} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho - 2s^{\sigma\tau} T_{\tau\rho}{}^\rho{}_{,\sigma} - \gamma^{\beta\tau} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho + \gamma^{\mu\nu} T_{\rho\nu}{}^\sigma T_{\mu\sigma}{}^\rho. \quad (22)$$

In order to make use of (7), we must multiply both sides of the equation (15) by  $\beta$ .

\* 'SitzBer. Preuss. Akad. Wiss.', vol. 11, p. 109 (1932).

We then obtain on the right-hand side

$$2\alpha^\sigma (\alpha^\tau T_{\tau\rho}{}^\rho \psi)_{;\sigma} - 2s^{\sigma\tau} T_{\tau\rho}{}^\rho{}_{;\sigma} \psi - 2\alpha^\sigma \alpha^\tau T_{\tau\rho}{}^\rho (\partial\psi/\partial x^\sigma) + 2s^{\sigma\beta} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho \psi \\ - \gamma^{\beta\tau} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho \psi + \gamma^{\mu\nu} T_{\mu\nu}{}^\nu T_{\mu\sigma}{}^\sigma \psi. \quad (23)$$

We can now substitute for  $\alpha^\tau T_{\tau\rho}{}^\rho \psi$  by (7).

The first term of (23) becomes

$$2\alpha^\sigma (\alpha^\tau \frac{\partial\psi}{\partial x^\tau})_{;\sigma} = 2\alpha^\sigma \alpha^\tau (\frac{\partial\psi}{\partial x^\tau})_{;\sigma} + 2T_{\beta\sigma}{}^\sigma \alpha^\sigma \alpha^\beta \frac{\partial\psi}{\partial x^\tau},$$

Since  $(\partial\psi/\partial x^\tau)_{;\sigma}$  is symmetric in  $\sigma$  and  $\tau$ , the first of the terms on the right becomes  $2\gamma^{\sigma\tau} (\partial\psi/\partial x^\tau)_{;\sigma}$ .

In the familiar form

$$\gamma^{\sigma\tau} \left( \frac{\partial^2\psi}{\partial x^\sigma \partial x^\tau} - \Gamma_{\tau\sigma}{}^\beta \frac{\partial\psi}{\partial x^\beta} \right)$$

we recognize the generalized Laplacian, which, in five dimensions, is written  $\bigcirc \psi$ .

Thus by making the substitution in (23) we obtain from (15)

$$\bigcirc \psi - s^{\sigma\tau} T_{\tau\rho}{}^\rho{}_{;\sigma} \psi + \alpha^\sigma \alpha^\beta T_{\beta\sigma}{}^\sigma \frac{\partial\psi}{\partial x^\tau} - \alpha^\sigma \alpha^\beta T_{\beta\sigma}{}^\sigma \frac{\partial\psi}{\partial x^\sigma} + s^{\sigma\beta} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho \psi \\ - \frac{1}{2} \gamma^{\beta\tau} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho \psi + \frac{1}{2} \gamma^{\beta\tau} T_{\beta\sigma}{}^\sigma T_{\rho\tau}{}^\sigma \psi = \frac{1}{2} \gamma^{\beta\tau} (P_{\beta\tau} - R_{\beta\tau}) \psi. \quad (24)$$

The interpretation of  $\bigcirc \psi$  in Klein's theory, or in de Broglie's modification, of it, is given in the first paper. A suggested value of  $\gamma^{\beta\tau} P_{\beta\tau}$  was also given.  $\gamma^{\beta\tau} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho$  is a familiar form, and if we write  $T_{\beta\sigma}{}^\sigma = -(2\pi i/\hbar) \Pi_\beta$ , its value is equal to that of  $\gamma^{\beta\tau} P_{\beta\tau}$ . If we then write

$$\gamma^{\beta\tau} P_{\beta\tau} = \gamma^{\beta\tau} T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho = -\frac{4\pi^2}{\hbar^2} \left( m^2 c^2 + \frac{e^2}{\gamma_{55} \alpha^2} \right)$$

we obtain the following group of terms from (16)

$$\bigcirc \psi + \frac{4\pi^2}{\hbar^2} \left( m^2 c^2 + \frac{e^2}{\gamma_{55} \alpha^2} \right) \psi + \frac{1}{2} R \psi - s^{\sigma\tau} T_{\tau\rho}{}^\rho{}_{;\sigma} \psi.$$

We have stated that there is a doubt about the identification of  $T_{\tau\rho}{}^\rho$ , and this equation (24) does not limit us sufficiently to remove it, for it would always be possible to assume for  $\frac{1}{2} \gamma^{\beta\tau} (P_{\beta\tau} + T_{\beta\sigma}{}^\sigma T_{\tau\rho}{}^\rho)$  the value we have just ascribed to  $\gamma^{\beta\tau} P_{\beta\tau}$ . The last term in the group just written down becomes

$$\frac{2\pi i e}{\hbar c} s^{\sigma\tau} \frac{\partial \phi_\tau}{\partial x^\sigma} \psi,$$

if we make the identification directly from Dirac's equations.

Again making use of the anti-symmetric property of  $s^{\sigma\tau}$  we can write

$$s^{\sigma\tau} \frac{\partial \phi_r}{\partial x^\sigma} = \frac{1}{2} s^{\sigma\tau} \left( \frac{\partial \phi_r}{\partial x^\sigma} - \frac{\partial \phi_\sigma}{\partial x^r} \right) = \frac{1}{2} s^{\sigma\tau} F_{\sigma\tau},$$

where  $F_{\sigma\tau}$  is the component of the electromagnetic field.

If we write  $T_{r\rho} = -(2\pi i/\hbar) \Pi_r$ , we arrive at the same conclusion by making use of the relation

$$\frac{\partial p_r}{\partial x^\sigma} - \frac{\partial p_\sigma}{\partial x^r} = 0,$$

which is true for the component  $p$ , in mechanical applications.

If the momentum components did not enjoy this property, and the velocity were of the nature of that occurring in hydrodynamics, it might not be possible to get rid of the  $p$ 's so simply, for in hydrodynamics the curl of the velocity is twice the velocity of rotation, and unless the motion is irrotational, an extra term is contributed here, indicating an additional energy term of magnitude  $\frac{\hbar}{2\pi} \omega$ , where  $\text{curl } p = 2m\omega$ .

The value of  $R$  was discussed in the first paper.

Thus (24) may be written

$$\left\{ \square + \frac{4\pi^2}{\hbar^2} \left( m^2 c^2 + \frac{e^2}{\gamma_{55} \alpha^2} \right) + \frac{1}{2} R - \frac{2\pi i e}{\hbar c} F_{\sigma\tau} s^{\sigma\tau} \right\} \psi$$

= the extra terms of (24). (25)

It does not seem easy to interpret the extra terms physically, but the occurrence of the terms of the quantum theory is remarkable.

The method suggests a new approach to the subject and offers a relativistic basis to the theory.

It is true that the relativistic aspect appears as a background and that in the present state of the quantum theory, when operators detach themselves to come forward to play their parts on the stage of physics, they lose some of the harmony of their setting. But this may be due to our failure to appreciate their true nature.

In the account here given the method is not exhausted of its possibilities. For example, if we attempted to follow up a suggestion made earlier, we might adopt the vanishing of the divergence of a tensor  $A^{\alpha\nu}$  for the quantum equations. This would lead us to consider the value of the function  $\psi$ , which made  $\alpha^\mu \alpha^\nu \psi$  replace  $A^{\mu\nu}$ , and which preserved the Riemannian form of the divergence.

One of the tests of the method must be the introduction of the spin term in the form in which it occurs in (25).

The results obtained in Schroedinger's paper may be compared with those of this paragraph.

*Some Remarks on the Fifth Co-ordinate and on Rest Mass.*

Two reasons can be offered for introducing a fifth co-ordinate; firstly in order to include electromagnetic and gravitational phenomena in one relativistic scheme; secondly in order to provide a co-ordinate to play the part of a conjugate to  $mc$  in the quantum equations.

The requirements in each case are satisfied by simple conditions.

The method of Kaluza and Klein gave particular values to the  $\gamma_{\mu\nu}$ . To take a case as typical

$$\gamma_{mn} = g_{mn} + \gamma_{55}\alpha^2\phi_m\phi_n, \quad (26)$$

$m$  and  $n$  here having the values 1 to 4.

$\gamma_{55}$  is taken to be constant and  $\alpha$  is a universal constant. There are, of course, values for the other components  $\gamma_{\mu\nu}$ , but these need not be given in detail.

With these values it is well known that the track of an electron in the two fields is a geodesic and if  $\alpha = e/mc \sqrt{-\gamma_{55}}$  the geodesic is a null geodesic.

If  $d\sigma$  denote the five-dimensional element and  $ds$  the four-dimensional, it can be shown that

$$d\sigma^2 = ds^2 + d\theta^2, \quad (27)$$

where

$$d\theta = \gamma_{5\nu} dx^\nu / \sqrt{\gamma_{55}}.$$

When the track is a null geodesic  $d\sigma = 0$  and

$$ds^2 = -d\theta^2, \quad (28)$$

$d\theta$  is the projection of  $d\sigma$  on the fifth axis, and since

$$ds^2 = -c^2 d\tau^2, \quad (29)$$

we see that  $d\theta$  plays the same part as the proper time.

In Proca's discussion, to which reference has been made, we find that  $\sigma$  is the co-ordinate conjugate to  $-mc$ .

We thus see that in a five-dimensional continuum, in which null geodesics are the physically important tracks, Proca's arrangement of conjugates and momenta is just that which would be expected.

His set is as follows :—

$$\left. \begin{aligned} x_1, x_2, x_3, x_4 &= \mathcal{U}, & x_5 &= c\tau \\ p_1, p_2, p_3, p_4 &= \mathcal{W}/c, & p_5 &= -mc \end{aligned} \right\}$$

In the theory of relativity

$$p_n p^n = m^2 c^2 \quad (30)$$

or in the special theory

$$\Sigma p_n^2 = -m^2 c^2. \quad (30A)$$

The corresponding five-dimensional formula is

$$\Pi_\mu \Pi^\mu = I^2, \quad (31)$$

where  $I^2 = m^2 c^2 + (e^2 / \alpha^2 \gamma_{55})$ , if the values of the Kaluza-Klein theory are adopted.

When  $\alpha = e/mc \sqrt{-\gamma_{55}}$ ,  $I^2$  vanishes and we return to the form (30). In the absence of an electromagnetic field (30) and (31) become identical.

Thus it appears that, both in the quantum theory and in the theory of relativity, we are concerned only with the null geodesics of the five-dimensional continuum. It is this fact that gives to  $x^5$  and to  $mc$  the special places they occupy in the theory and which disturbs the general relativistic character of the quantum equations.

We may say that the appropriate relativistic theory is attained by the use of a five-dimensional continuum, but that we make use only of the null geodesics of that continuum.

It is not unreasonable to extend the relation (30) and write in its place (31), where  $I$  does not in general vanish. It does so only for certain tracks of space. These tracks are the ray tracks of the continuum corresponding to the light ray tracks of relativity. This analogy brings out what may be the real significance of particle waves.

If  $I$  does not vanish then the momentum components,  $\Pi_\mu$ , play a part like that of the  $p_n$  in the four-dimensional continuum.

In the conservation of the momentum,  $p_n$ , we write the two principles of conservation of energy and of momentum. We may follow the obvious suggestion that the conservation of  $\Pi_\mu$  corresponds to these principles and to a further principle.

If we examine the special value of  $\Pi_5$  on a null geodesic the suggestion is that the rest mass must be conserved.

This is of interest in connection with suggestions that radiation may arise from the union of such particles as protons and electrons or positrons and electrons.

Before the discovery of the positron it was suggested that radiation might arise from the union of protons and electrons.

This would mean that two particles of differing rest mass united to produce a photon of zero rest mass.

In a previous paper\* this was discussed and it was pointed out that the components,  $\Pi_s$ , for these particles appeared in the theory to have opposite signs and that they could unite to give a particle with  $\Pi_s = 0$ , provided there was an asymmetry of scale in association with the two bodies. The proton appeared as a diminished electron, in which values were scaled down, and the small units of the proton were taken to be equal to the larger units of the electron.

The discovery of the positron makes it possible to remove the metric asymmetry and the suggestion is that a positron and electron unite to give a particle with  $\Pi_s = 0$ , *i.e.*, to give a photon, while we should expect this to be impossible in the case of proton and electron.†

The theory will permit us to describe the rest mass of one particle as positive and of the other as negative, but it is difficult to attach a physical meaning to this. There is doubtless some phenomenon which corresponds to what we describe as negative energy and possibly also to what we describe as negative rest mass, though these terms may not be happily used in this connection.

### *Summary.*

This is a second contribution under this title and the quantum equations are obtained in a way which throws a new light upon them and fits them into a consistent scheme.

The second order equation is derived, not by an operational process, but by using the first order equations as a substitution in the equation of curvature. The result contains all the terms required in matrix mechanics in the description of the electron. It contains also additional terms for which no explanation is offered.

The treatment is a relativistic one and a discussion on the difficulties associated with the form of Dirac's equation is given.

\* Flint, 'Proc. Roy. Soc.,' A, vol. 131, p. 170 (1931).

† Cf. Furth, 'Z. Physik,' vol. 85, p. 294 (1933).

---



## *On the Polarization of Electrons by Scattering.—II.*

By E. G. DYMOND, M.A., Carnegie Teaching Fellow in the University of Edinburgh.

(Communicated by C. G. Barkla, F.R.S.—Received February 8, 1934.)

In a previous paper\* experiments were described to detect a polarization of a beam of electrons which had suffered large angle deflection from heavy nuclei. The work was carried out to check the theoretical conclusions of Mott,† and the experimental conditions followed those laid down by him as closely as possible.

A beam of electrons of homogeneous velocity fell on a gold film, sufficiently thin to ensure single scattering, from which they were scattered through  $90^\circ$  to a second film and again scattered through  $90^\circ$  to a Faraday cylinder connected to an electrometer. The second film and Faraday cylinder could rotate about an axis joining the two films, so that the direction of the final beam was either parallel or antiparallel to its original direction.

The results of Mott's calculations are shown in fig. 1, in which the asymmetry in the beam (expressed as a percentage difference in the two scattering directions parallel and antiparallel to the original beam) after  $90^\circ$  deflections from single nuclei of gold is plotted against the energy of the electrons.

The work described in I was carried out with a maximum voltage of 70 kv., for which the theoretical value of  $2\delta$ , the asymmetry, is about 10%. The experimental value found was  $1.7 \pm 0.3\%$ , where the probable error was calculated from the internal consistency of the measurements. This discrepancy called for further investigation, and it seemed desirable to go to higher voltages for the sake of the increased effect to be expected from a study of Mott's curve.

### *Apparatus.*

Advantage was taken of the experience gained largely to rebuild the apparatus; the tube for supplying the electron beam was completely redesigned to stand the higher potentials, and at the same time a much more rigid construction was adopted, so that fluctuations in the scattered current due to slight movements of the filament relative to the defining apertures were considerably

\* 'Proc. Roy. Soc.,' A, vol. 130, p. 638 (1932), hereafter cited as I.

† 'Proc. Roy. Soc.,' A, vol. 124, p. 425 (1929), and vol. 135, p. 429 (1931).

reduced. Thermal expansion of the various parts, due either to heat from the filament or to that evolved at the anode by electron bombardment, was sufficient to cause a steady change in the scattered current. In the latter part of the work a Faraday cylinder was introduced to catch part of the main beam after it had traversed the first scattering foil. This electrode measured the current scattered from the foil within a cone of semiangle  $5^\circ$ ; for a given voltage this current was proportional to the incident beam, and could accordingly be used as a control.

The various defining apertures in the beam were all reduced in diameter as much as was consistent with intensity. While the outer frame of the

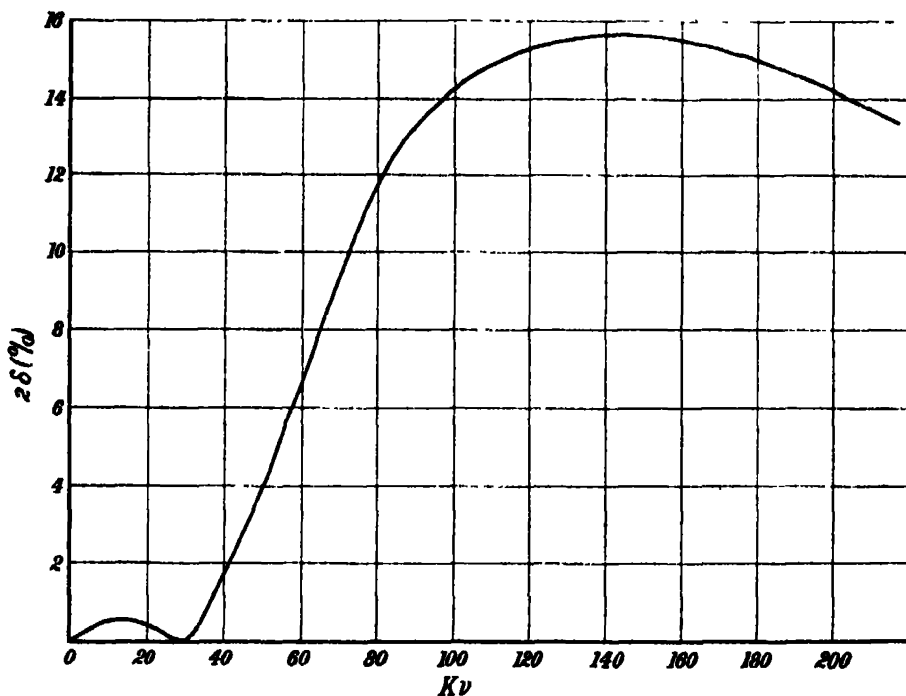


FIG. 1.—Mott's curve for the variation of polarization with voltage.

apparatus was the same as that used in I, all the internal parts were renewed, so that it is unlikely that any asymmetry introduced into the scattering by the geometrical conditions persisted from the old apparatus to the new.

The potential source for accelerating the electrons was a voltage doubling circuit of the type developed by Cockcroft and Walton,\* fed from a transformer giving 100 kv. peak. By this means steady potentials up to 200 kv. could be obtained with a ripple of 2% with the outputs used. The potential on the

\* 'Proc. Roy. Soc.,' A, vol. 136, p. 619 (1932).

tube was measured by a generating voltmeter, such as is described by Kirkpatrick,\* with minor changes to suit the different operating conditions. The voltmeter was calibrated against a sphere spark gap. The readings were liable to be in error by several per cent. in absolute magnitude, but the chief function of the voltmeter was to enable the potential to be held steady within narrow limits during the course of observations. This was very necessary as the double scattered current varies approximately as  $V^{3.4}$ . Supply voltage variations are the principal cause of fluctuations in the readings, consequently automatic control would be an advantage.

It was shown in I that the influence of X-rays produced either at the first defining aperture or elsewhere was negligible up to a potential of 70 kv. With the new apparatus this was no longer so, and at 160 kv. large deflections of the electrometer were produced even with no scattering films in the apparatus. The effect was largely due to ionization of the air surrounding the electrometer and its lead (which were, of course, encased in earthed shields). As the leak due to X-rays varied with the azimuth of the collecting electrode a spurious asymmetry was introduced which might be mistaken for a genuine polarization effect.

It was not possible for mechanical reasons to absorb the X-rays completely in lead screens, as the weight required was prohibitive. Accordingly it was decided to evacuate the electrometer and lead system, to remove all disturbances from this cause. The electrometer was housed in a brass chamber, which was fitted with a small glass window to admit light to the mirror, and with external controls for adjusting the zero and sensitivity of the instrument. A short tube fitted with a Tombac flexible joint and ground cone joint to allow rotation of the Faraday cylinder connected the electrometer chamber to the scattering apparatus, so that no part of the insulated system was surrounded by air.

A high vacuum around the electrometer was not found desirable as not only did the instrument become highly undamped, but also the radiometer effect due to the heating of the mirror by the illumination became very marked. At a pressure of about 0.5 mm. of mercury the electrometer was critically damped and showed no radiometer effect.

The effect of evacuation was very definite; not only was the X-ray drift, determined by removing the scattering foils, almost entirely eliminated, but also the natural leak of the instrument was reduced, presumably owing to the

\* 'Rev. Sci. Instr.' vol. 3, pp. 1, 430 (1932).

greater dryness of the quartz insulation, and the general steadiness was very much improved.

The scattering foils used in I were commercial beaten gold of about  $9 \times 10^{-6}$  cm. thickness. As mentioned there, these foils are very inhomogeneous and it was suggested that plural scattering from the thicker portions (which are opaque when examined under a low power microscope) might be a cause of the small amount of polarization found, as it is not expected that plural scattered electrons should show polarization effects. Accordingly the foils for the present work were prepared by cathodic sputtering on a cellulose acetate base, which was subsequently dissolved away. This gave a metal film of great uniformity, in which no inhomogeneities could be seen with a low power, and it allowed of the preparation of films of different thickness. The thickest film was  $2.5 \times 10^{-5}$  cm., while the thinnest was probably 4 to  $5 \times 10^{-6}$  cm., although it was too thin to measure by weighing.

It may be said here that the foil thickness has no influence on the results. Parallel experiments first with beaten and then with sputtered films of the same thickness give the same results, which agree with those found in I. Further, direct measurement of the variation of scattered intensity with voltage gives strong evidence that single scattering takes place at all potentials above 80 kv. This point will be discussed more fully below.

Some of the foils through long use became opaque where the electron beam had struck them, presumably caused by the evaporation of tungsten from the filament. Their behaviour in producing an asymmetry was the same in this condition as when first introduced into the apparatus.

In order to prevent slow electrons and positive ions from the residual gas from entering the Faraday cylinder the same procedure was adopted as in I of placing two aluminium foils of combined thickness  $1.4 \times 10^{-4}$  cm. in front of the collector. Subsidiary experiments showed that of electrons with energy 10 kv. only 1% reached the cylinder, while of those of 90 kv. the number was 75 to 80%, expressed in terms of those arriving at 160 kv. Such an aluminium filter does not exclude all electrons which have suffered inelastic collisions on the way, but it does exclude completely the most numerous class of slow secondaries.

### *Results.*

The method of taking and combining observations was the same as that employed in I. In order to eliminate as far as possible purely geometrical asymmetries, it was assumed that at low voltages there is no polarization

effect, and that any asymmetry found at these voltages was of instrumental origin. This instrumental asymmetry was always subtracted from that observed at higher voltages to arrive at the supposed true polarization effect. The low voltage used was a reference zero for  $\delta$  varied between 15 and 30 kv., depending on the thickness of the foils used. In the course of the work it was found that this proceeding, which has also been adopted by other workers in this field, gave rise to fallacious results, unless combined with other methods of control.

The first point to be attacked was the question of the influence of the inhomogeneity in thickness of the scattering foils as outlined above. A series of measurements was carried out first with beaten, and then with sputtered films of somewhat less thickness. The values obtained at 90 kv. for  $\delta^*$  were  $+1.1 \pm 0.7\%$  for the beaten, and  $+1.0 \pm 0.3\%$  for the sputtered foils. The method of production of the foils, therefore, has no influence on the results. It is noteworthy that the value of  $\delta$  found with the new apparatus is actually slightly less than that previously determined ( $1.7 \pm 0.3\%$ ) at 70 kv. The difference though hardly significant at once leads to the suspicion that the observed effect is not a true polarization effect at all. Reference to Mott's curve, fig. 1, shows that  $\delta$  rises most steeply in this region of voltage.

A sharp rise in the value of  $\delta$  was originally found at higher voltages than 90 kv., accompanied by greater fluctuations in the values. This was shown to be largely due to X-rays affecting the electrometer and the lead system, as previously mentioned. The rotating part of the apparatus containing the collecting electrode and second scattering foil screened the electrometer, and as it is moved the X-ray leak in the electrometer changed, in the required sense to account for the observed effect. Blank experiments with no scattering films showed that this leak began to be important at about 120 kv.

The electrometer and lead were then evacuated resulting in almost complete elimination of the X-ray effect; also the whole apparatus was taken down and reassembled.

The results thus obtained are shown in Table I, group A. It was found that  $\delta$  had changed sign and was increasing in the opposite direction with increasing voltage. Such a result was very surprising and as it was suspected that bad adjustment might be the cause, the whole apparatus was again taken apart. It was, in fact, found that in the previous reassembly the first scattering film had been replaced somewhat too high, so that the intersection of the original

\* The sign of  $\delta$  is reckoned positive when the antiparallel is stronger than the parallel scattering.  $\delta$  is theoretically always positive.

Table I.

Group.	kv.	$2\delta$ .
A	90	$-5.9 \pm 2.3$
	125	$-7.8 \pm 2.4$
	160	$-9.6 \pm 4.1$
B	125	$+2.7 \pm 0.6$
	160	$+4.1 \pm 0.2$
C	90	$-1.0 \pm 0.6$
	125	$-0.7 \pm 0.8$
	160	$+0.4 \pm 0.6$

electron beam with the film was no longer on the axis of rotation of the second film and collector. This introduced a considerable geometrical asymmetry into the scattering. It had always been assumed previously that any such asymmetry would be independent of electron velocity, and so would be eliminated by the method of combining observations at a high and a low voltage. It seems from these experiments that this is not always a justifiable procedure.

In order to test this conclusion, the error in position of the first film was somewhat over corrected. The value of  $\delta$  now changed sign, as shown in the group B. On readjusting the height of the film as accurately as possible the values in group C were obtained.

It is therefore apparent that by accurate adjustment the value of  $\delta$  can be reduced to zero, and by deliberate maladjustment may be made to assume either positive or negative values. There is no evidence of a true polarization larger than 1% in these measurements, and there is no reason to believe that the small positive result,  $2\delta = 1.7\%$  at 70 kv., found in I is not due to a slight error of adjustment as just described.

#### *Discussion.*

Of the various causes likely to reduce the observed values of  $\delta$  below the theoretical values the most important is the stray scattering from parts of the apparatus other than the films themselves. This can be investigated, as described more fully in I, by removing successively the two films. This stray scattering does not decrease with increasing velocity nearly so fast as the single nuclear scattering, and consequently becomes relatively more important at the higher energies. The X-ray effect also is measurable, though small, at 160 kv. The contribution of these two effects to the measured current, in the last series of determinations, group C, amounted to 1.5% at 30 kv., 4.5% at 125 kv., and 7.5% at 160 kv. The stray scattering, although small, was

highly unsymmetrical, the asymmetry decreasing with increasing energy. It is highly probable, though not certain, that the asymmetries set out in Table I are due to the stray effect. It is not possible to show this accurately, as the stray scattering is not quite the same when the films are in position as when they are withdrawn. There is always a contribution to the stray effect, for which a correction cannot be made, from electrons scattered from the second film to the walls and back to the collector. Nevertheless, the order of magnitude and sign of the effect is sufficient to justify the conclusion that the observed asymmetry has its origin here.

Much more weighty evidence for this conclusion is the fact that the observed value of  $\delta$  can be controlled both as regards sign and magnitude by small adjustments (of about 2 mm. in all) in the height of the first scatterer with respect to the axis of rotation. Once this fact is established it is impossible to accept the existence of a true polarization effect exceeding 1%. Whether an effect of less than this exists the experiments are not accurate enough to decide, and it would need radical alteration in design to increase the accuracy sufficiently.

The origin of the observed variations with energy in the value of  $\delta$  and their dependence on the position of the first film is somewhat obscure. In order to account for the facts it is necessary that the ratio of the number of electrons received, parallel and antiparallel to their original direction, should depend on their velocity. It is relevant here to notice that in order to obtain sufficient intensities at the higher velocities rather thick films must be used. It is probable that single scattering takes place above 80 kv., but it is certain that it does not at 30 kv. The distribution in angle of the electrons is widely different in the two cases, being much more uniform at the lower velocity.

As the angular widths of the beams scattered from the two films are finite the mean angle of scattering will be different in the two azimuths, unless the axis of the beam coincides exactly with the axis of rotation. It is therefore possible to explain qualitatively the experimental observations, although their magnitude is somewhat surprising. But whatever their explanation, their existence is not in doubt and it invalidates the previous assumption of the author, and other workers, that an asymmetry which varies with velocity of the electrons is necessarily due to a polarization effect.

The failure of this experiment to agree with theory might be due to one or more of the following factors :—

- (a) Inelastic scattering at the films, and radiation braking.
- (b) Stray scattering not experimentally accounted for.

(c) Plural scattering.

(d) Nuclear screening, invalidating Mott's assumption of a pure Coulomb field.

Electrons scattered by the processes *a*, *b*, and *c* cannot be expected to show any polarization; only the fraction elastically and singly scattered should contribute, and if this fraction is small the expected polarization must be small. With regard to *d*, if the electrons do not move in a Coulomb field the theory is not applicable and at present it is impossible to say what polarization is to be expected. The influence of the first two factors *a* and *b* cannot be exactly determined without having recourse to magnetic analysis of the beam after the second scattering. It has been stated that the aluminium filter in front of the collector excludes all electrons of less than 10 kv. energy, but that its action is only partial for faster electrons. However, it is well known that a very high proportion of the secondary electrons have very small energies. Taking this into account together with the fact that the total number of secondaries as experimentally determined is small, it is safe to presume that only a small fraction of the total measured current can be due to secondaries unaccounted for by direct determination.

With regard to those electrons actually scattered from the films, as long as single scattering takes place energy losses can only arise from excitation or ionization of the deeper levels of the gold atoms or by excitation of the continuous X-ray spectrum. Such inelastic impacts will result in electrons which can penetrate the aluminium filter, although in reduced numbers. There is no experimental work on the absolute probability of ionization of the X-ray levels. The theory of Bethe,\* however, will give us the right order of magnitude. He finds that for large angles of scattering the ratio of the inelastic to the elastic scattering is  $1 : Z$ , i.e.,  $1 : 79$ . This will not be true in the immediate neighbourhood of the K ionization potential (80,000 volts), but none the less probably represents the order of magnitude. It is quite justifiable therefore to neglect the inelastic scattering. It should be noted that a very large amount of inelastic scattering is required to bridge the gap between the experimental result (less than 2%) and the theoretical (15%).

The contribution of the continuous X-ray spectrum to the inelastic scattering is known to be of the same order of magnitude as that of the line spectrum.† Further, Mott‡ has discussed the influence of the continuous spectrum in

\* 'Ann. Physik,' vol. 5, p. 325 (1930).

† Hansen and Stoddard, 'Phys. Rev.,' vol. 43, p. 701 (1933).

‡ 'Proc. Camb. Phil. Soc.,' vol. 27, p. 265 (1931).



scattering experiments and found that it amounted at most to 2 or 3%, thus confirming the indirect experimental evidence.\* We may therefore conclude that the measured current is predominantly due to elastic scattering.

The third factor, that the films may not be thin enough to ensure single scattering, can be investigated by comparing the observed variation with voltage of intensity of the final beam with the theoretical values calculated by Mott.† He finds that the intensity is given by

$$I = \left[ \frac{Ze^2}{2mv^2} \operatorname{cosec}^2 \frac{\theta}{2} \right]^4 \left[ 1 - \left( \frac{V}{c} \right)^2 \right]^2 R^2,$$

where  $\theta$  is the angle of scattering, here  $90^\circ$ ,  $v$  the velocity of the electron, and  $R$  a numerical factor varying with  $v$ , whose order of magnitude is unity and for which he tabulates values.

Before comparing the experimental observations with this expression certain corrections must be made, the most important of which is that due to the absorption of the aluminium filter in front of the collector. This can be determined by independent experiment. The correction is small for the higher velocities, but becomes large and therefore unreliable below 50 kv. Further quite small corrections for stray scattering and X-ray effect on the electrometer were also made. The result is shown in fig. 2, where the intensity of double scattering against voltage is plotted logarithmically. The theoretical and experimental curves are fitted at 100 kv.

It is seen that above 80 kv. they agree closely, the agreement becoming progressively worse as the voltage is lowered. Owing to the corrections too much reliance must not be placed on the experimental result, but it tends to confirm the assumption that single scattering is taking place, and that Mott's results for the polarization should be applicable.

With regard to the last point—the screening of the nucleus—a 130-kv. electron scattered classically at  $90^\circ$  will not penetrate within the K ring of gold, so that the atomic electrons might be expected to influence the polarization unfavourably. Sauter,‡ however, has discussed this point and finds that screening has no perceptible effect except for slow electrons, when no polarization is expected.

\* See also a paper by Sauter, 'Ann. Physik,' vol. 18, p. 486 (1933), who comes to the conclusion that the energy loss of an electron through radiation is small compared with that through ionization.

† 'Proc. Roy. Soc.,' A, vol. 135, p. 429 (1932).

‡ 'Ann. Physik,' vol. 18, p. 61 (1933).

We are therefore driven to the conclusion that the theoretical results are wrong. There is no reason to believe that the work of Mott is incorrect\* ; moreover, Sauter, in the paper last quoted, has attacked the same problem from a different angle, using the Born collision method, and obtains a result identical with Mott's, to the approximation of Mott's first paper. Unfortunately this is not adequate for heavy nuclei and Sauter has not yet carried the calculations to a further degree of approximation.

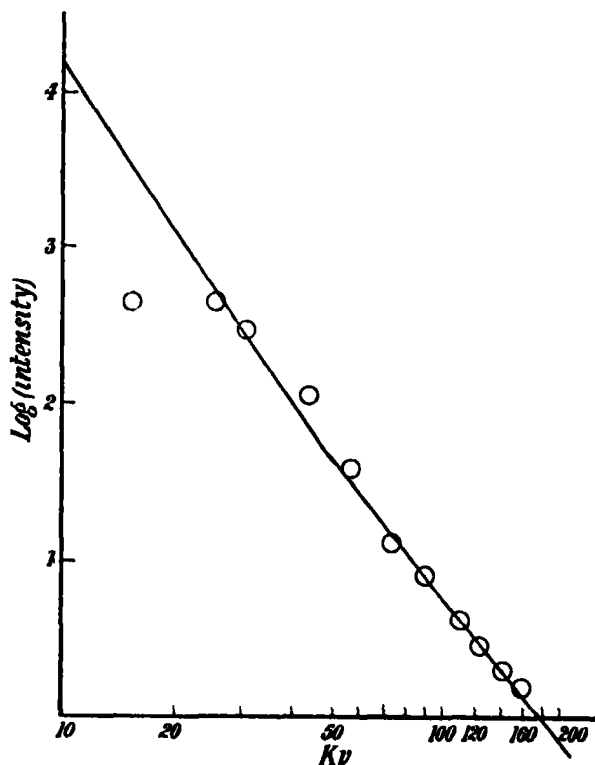


FIG. 2.—Variation of double scattered intensity with voltage. Full line, theoretical curve of Mott ; circles, experimental points. The experimental curve has been fitted to the theoretical at 100 kv.

It seems not improbable therefore that the divergence of theory from experiment has a more deep-seated cause, and that the Dirac wave equation needs modification in order to account successfully for the absence of polarization.

Rupp† has carried out experiments similar to those described here, in which he finds a positive effect of 4% at 130 kv. and 10% at 250 kv. His first

\* Professor Mott has kindly informed me that the results of his second paper have recently been rechecked carefully.

† 'Z. Physik,' vol. 79, p. 642 (1932).

scatterer was some 10 times thicker than that used in this work, while the second scatterer was a wire of 0.12 or 0.2 mm. diameter. Multiple and non-elastic scattering in his experiments must play a much more important role. Slow electrons were hindered from reaching the collecting electrodes by gold foils (thickness not given), in the same manner as described here.

Rupp finds that a transverse magnetic field between the two scatterers of 440 gauss, accompanied by an electric field to bring the deflection of the beam to zero, reduces the asymmetry to nothing. Such behaviour is to be expected if the beam is made up of electrons whose axes of spin are predominantly in one direction. But it is not clear that the inverse conclusion, which he draws, that a polarization of the beam is the only possible cause for this effect, is justified. An asymmetry of purely instrumental origin may also be influenced by the field in this manner. G. P. Thomson\* has recently repeated some earlier experiments of Rupp in which the second scatterer was a thin film normal to the beam. Thomson failed to confirm the observation of Rupp that the diffraction rings formed by the second foil have an asymmetric intensity distribution round their circumference. In view of this lack of confirmation it is difficult to accept Rupp's results in this experiment as due to a real polarization; nevertheless, his asymmetry disappeared in a magnetic field in the same way as that found in his later work, just described. Finally it should be recalled that the author has found asymmetries of the same order of magnitude and even larger than those reported by Rupp; but subsequent work showed that these effects were spurious.

This work was made possible by Grants from the Earl of Moray Endowment, and from the Ritchie Fund of the University of Edinburgh to defray the cost of the high tension transformer, and from the Government Grant Committee of the Royal Society for the voltage multiplying equipment.

### *Summary.*

The author's previous experiments have been extended to higher electron velocities with an improved apparatus.

No asymmetry in the scattering which could be ascribed to a polarization effect greater than 1% has been observed, up to an energy of 160 kv., whereas Mott predicts a maximum asymmetry of 15%. The effect previously observed of 1.7% at 70 kv. has been shown not to be real.

\* 'Nature,' vol. 132, p. 1006 (1933).

The influence of stray, inelastic, and plural scattering, and of radiation braking on the results are discussed and are dismissed as being inadequate to account for the lack of agreement.

In view of Sauter's confirmation, by another method, of Mott's work, it would seem that modification in the relativistic wave equation is required to account for these results.

---

### *On the Fracture of Fibred Iron-Silicon Sheets.*

By FRANCIS BITTER, John Simon Guggenheim Memorial Fellow.

(Communicated by G. I. Taylor, F.R.S.—Received February 16, 1934.)

[PLATE 17.]

Flat strips made of material that is neither too soft nor too brittle have a tendency to break at a specific angle.\* A few experiments have been carried out on sheets having a particularly symmetrical fibre-structure in order to discover whether any relationship exists between grain orientation and fracture in this material. The results indicate that such a relationship exists, and that further investigations on the mechanical properties of fibred materials may be helpful in developing a satisfactory interpretation of the strength of solids.

#### *The Material Investigated.*

The substance from which the sheets under discussion were made was especially prepared electrolytic iron plus 4% silicon melted in hydrogen. This has a body-centred cubic lattice structure, and when rolled by the usual methods has a fibre-structure with a (100) plane in the surface of the sheet, and a [110] axis in the rolling direction. The chief departure from the above consists in a considerable scattering of orientations about the axis in the rolling direction. By means of a special treatment, which it is intended to report elsewhere, it is possible to reduce or to remove this asymmetry, so that sheets having identical fibring about axes in the plane of the sheet parallel and perpendicular to the rolling direction are obtained.

\* Körber and Hoff, 'Mitt. K.Wilh.-Inst. Eisenforsch.,' vol. 10, p. 175 (1928); Körber and Siebl, *ibid.*, vol. 10, p. 189 (1928).





FIG. 2.



FIG. 3.



FIG. 4.

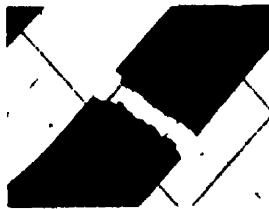


FIG. 5.

FIGS. 2-5. —Fractures obtained in strips cut from a sheet at angles of 0°, 15°, 30°, and 45° to the rolling direction and immersed in liquid air.

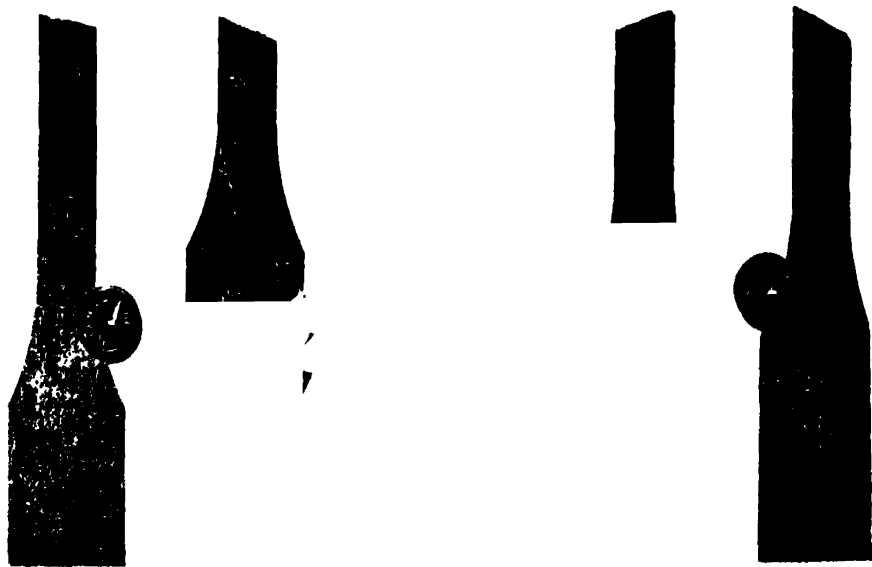


FIG. 6. —Fractures obtained by strips cut as in figs. 2-5, but pulled at about 200° C.

Two such sheets were investigated, and for each experiment samples from each of the sheets were used. Strips were cut out of the sheets in several directions, making angles of  $0^\circ$ ,  $15^\circ$ ,  $30^\circ$ , and  $45^\circ$  with the rolling direction.

*Fractures at Liquid Air Temperature.*

The first experiments were carried out at the temperature of liquid air with the help of the Dewar tube shown in fig. 1. The outside of the Dewar tube was corrugated in order to prevent its breaking, through the contraction of the inner wall when filled with liquid air. The outside of the cork and the slot through which the specimen passed were smeared with vaseline to prevent the liquid air from leaking from the container. The ends of the specimen were clamped into the grips of a 10-ton Amsler machine capable of exerting tension, and the specimen was pulled until it broke. No plastic deformation was observed, and whatever there was, was small compared to the elastic extension.

Typical photographs of the fractures are shown in figs. 2, 3, 4, and 5, Plate 17. The samples were mounted on graph paper in such a way that the diagonals of the squares are parallel and perpendicular to the rolling direction, and therefore parallel to  $[110]$  axes. Consequently, the lines on the graph paper are parallel to the tetragonal or (100) planes of the crystal grains in the sample. In figs. 2, 3, 4, and 5 the samples were cut at angles of  $0^\circ$ ,  $15^\circ$ ,  $30^\circ$ , and  $45^\circ$  to the rolling direction. The point here illustrated is that in the above brittle fractures there is a marked tendency for break to occur parallel to the lines on the graph paper (or (100) planes) independently of the orientation of the applied tension. In figs. 2 and 5, the fracture consists of smaller zig-zags than in figs. 3 and 4. In figs. 3 and 4 it is especially apparent that the two sides of the fracture no longer fit each other, and that small pieces must have been broken from the ends of the sample. There is a marked difference in the tension necessary to break the samples cut in various directions from the sheet, as shown in Table I.

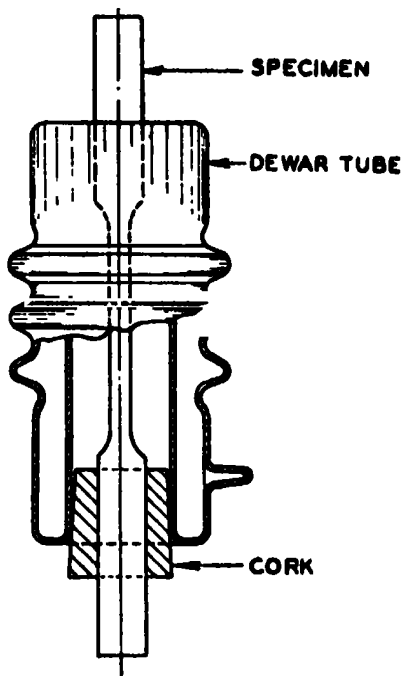


FIG. 1.—Diagram of apparatus for tensile tests in liquid air.

It is interesting to observe that least force is required to break the sample when the tension is normal to the (100) planes along which the break takes place. It may also be significant that the forces in the first and last samples in Table I have a ratio of 1.4 or about  $\sqrt{2} = 1/\cos 45^\circ$ , the component of the applied tension normal to a (100) plane being proportional to  $1/\cos \theta$ .

Table I.—Breaking Strength of Fibred Samples at Liquid Air Temperature.

Angle between axis of sample and rolling direction.	Force required to break sample, lbs.
°	
0	540
15	490
30	435
45	390

#### *Fractures at Higher Temperatures.*

Experiments similar to those described in the previous paragraph were performed at temperatures ranging from room temperature to about 500° C. In the neighbourhood of 200° C. breaks having straight edges making an angle between 55° and 80° with the axis of the sample were obtained. In the higher ranges of temperature, there was a great deal of plastic deformation and the breaks had jagged edges. At the lower temperatures, the breaks often were made up of two parts, one parallel to (100) planes and showing the characteristics of brittle fractures, the other part of the fractured edge being at an angle, and similar to the break at 200° C. Samples showing typical smooth breaks obtained at 200° C. are shown in fig. 6, Plate 17. On the left is one of 12 samples cut parallel to the rolling direction. Nine of these broke at an angle of 75°, the other three did not have breaks with smooth edges. Next to this in fig. 6 are two samples cut out of the same sheet at an angle of 15° to a rolling direction. It will be observed that the two angles are not equal. They are 68° or 79°, depending on whether the acute angle is on the left or right of the sample. The situation is illustrated in fig. 7. The co-ordinate axes may be thought of as drawn on a test sample parallel to the tetragonal axes of the crystals in the fibre structure (at 45° to the rolling direction). The length of the sample is in the direction  $\sigma$  in which tension is applied. Break occurs along one of the two lines shown in the figure. These lines do not make equal angles with  $\sigma$ . Fourteen samples were cut at an angle of 15° to the rolling direction, of which nine broke with the acute angle on the left when viewed as in fig. 6. Of these nine,



seven broke at an angle of  $68^\circ$  as shown in fig. 6, and two did not show smooth edges. The remaining five samples cut at  $15^\circ$  to the rolling direction broke with acute angles on the right, three at an angle of  $79^\circ$ , and two at substantially different angles. These two samples fall completely out of the picture to be described below, and as they are the only two exceptions among the 51 samples examined they will be neglected. Further investigation is needed to ensure that their occurrence was due to some such accidental cause as local imperfections in the fibring of the material. Twelve samples were cut at an angle of  $30^\circ$  to a rolling direction, of which 10 had an acute angle on the left of  $58^\circ$  and two had an acute angle on the right of  $71^\circ$ . Thirteen samples were cut at an angle of  $45^\circ$  to the rolling direction, of which all broke at an angle of  $60^\circ$ , whether

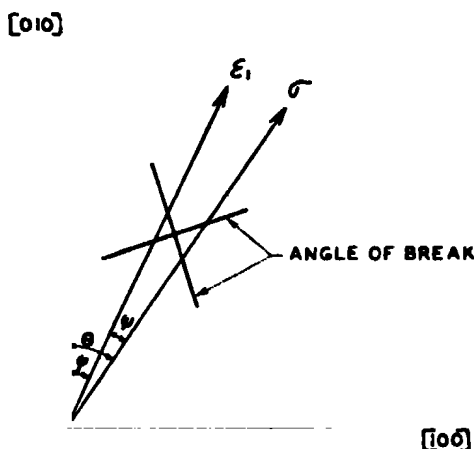


FIG. 7.—Diagram showing the designation of the angles defining maximum elongation  $\epsilon_1$ , and maximum stress  $\sigma$ , with respect to crystallographic axes in a fibred sheet.

the acute angle was on the left or right of the sample. In order to complete the description of the samples shown in fig. 6, it is necessary to say that all samples are shown in the photographs with the same side of the sheet, from which they were cut, towards the observer, and that, in their original positions in the sheet, the angle from the strip on the left to that on the right is  $45^\circ$  in a clockwise direction.

#### Discussion.

The observational material obtained on the angle of break has been reported in the last paragraph. The origin of the phenomenon observed is not known, but a few further comments may be of interest. Iron crystals are not elastically isotropic, and sheets in which certain grain orientations are predominant are

therefore not isotropic either. It is a well-known fact that in non-isotropic materials, the direction of maximum elongation does not in general coincide with the direction in which the tensile force is applied. Thus if, in a crystal, we examine a volume element which, in the undistorted state was spherical, we find that after the application of tension, the volume element has been drawn out from a sphere into an ellipsoid, but that the major axes of the ellipsoid are not parallel and perpendicular to the applied tension. Assuming that to a first approximation our fibred sheet has the elastic symmetry of a cubic crystal, and placing the cubic axes  $i, j, k$  parallel to the tetragonal or cubic axes of the crystal grains,  $i$  and  $j$  being in the plane of the sheet and making angles of  $45^\circ$  with the rolling directions, and  $k$  perpendicular to the plane of the sheet, we have for the tensor components of the elongation for any tensile stress in the  $i$ - $j$  plane :

$$A_{ii} = (S_{11} - S_{12}) F_{ii} + S_{12} (F_{ii} + F_{jj})$$

$$A_{ij} = S_{44}/2 F_{ij},$$

where  $S_{11}$ ,  $S_{12}$ , and  $S_{44}$  are the three elastic constants required to describe the elastic properties of a cubically symmetrical structure. Now, if  $\theta$  and  $\phi$  determine the directions of the applied force and of the maximum elongation respectively, and  $\psi$  is the angle between these directions, as shown in fig. 7, it can be shown that, writing

$$\wp = 2 (S_{11} - S_{12})/S_{44},$$

$$\wp \tan 2\phi = \tan 2\theta. \quad (1)$$

For isotropic material  $\wp = 1$ , and, therefore,  $\theta = \phi$  and  $\psi$  is always zero. The maximum elongation is parallel to the applied stress. Since  $\phi + \psi = \theta$ , it is possible to calculate  $\psi$  as a function of the direction in which forces are applied in the  $i$ - $j$  plane for various degrees of anisotropy, or various values of  $\wp$ . Plots of equation (1) are shown in fig. 8. The object of this calculation is to determine whether the angles of break as shown in fig. 7 are symmetrical with respect to the direction of maximum elongation. This was done by drawing a line bisecting the angle between the two directions of break (see fig. 7) and seeing whether the orientation of this line changed with  $\theta$  in the manner prescribed by equation (1). The points so observed fall within the shaded areas in fig. 8, the area covered by the shading being a rough estimate of the experimental inaccuracy in measuring the angles in question. The fact that the points do follow the trend of the curves indicates that the sug-

gestion is not impossible, and that a value of  $\varphi \sim 1.6$  is to be expected unless the agreement shown in fig. 8 is fortuitous.

The elongations for tensions parallel to a [100] and a [110] axis respectively are given by

$$\epsilon_{[100]} = S_{11}f, \quad (2)$$

$$\epsilon_{[110]} = \left( \frac{S_{11} + S_{12}}{2} + \frac{S_{44}}{4} \right) f, \quad (3)$$

so that by measuring the elongation produced by a given force acting on strips cut parallel to [100] and [110] axes, it is possible to measure two of the three

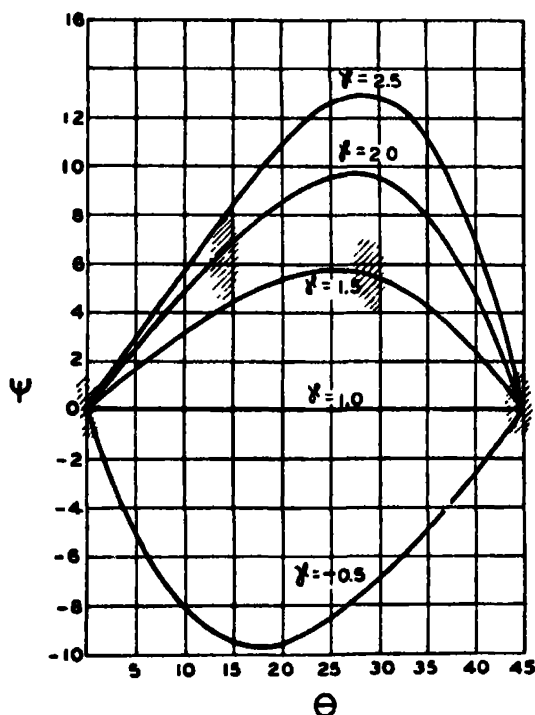


FIG. 8.—The relationship between the direction of maximum elongation and maximum stress in materials having different degrees of elastic anisotropy  $\varphi$ .

constants appearing in  $\varphi$ . In order to estimate the third, it is desirable to introduce the compressibility  $K$  where

$$1/3K = S_{11} + 2S_{12},$$

and to rewrite equation (3) in the form

$$\epsilon_{[110]} = \left[ S_{11} + \frac{1}{4} \left( 3S_{11} - \frac{1}{3K} \right) \left( \frac{1}{\varphi} - 1 \right) \right] f. \quad (4)$$

For iron  $3S_{11} = 22 \times 10^{-13}$  while  $1/3K$  varies between 1.7 and  $2.2 \times 10^{-13}$ . To a first approximation, we may, therefore, put  $1/3K = 0$  and obtain from equations (2) and (4)

$$\frac{1}{\rho} = \frac{1}{3} \left( 4 \frac{\epsilon_{[110]}}{\epsilon_{[100]}} - 1 \right), \quad (5)$$

where now  $\epsilon_{[110]}/\epsilon_{[100]}$  means the ratio of the elongations in the two directions produced by equal forces. This ratio was found to be 0.81 for the sheets in question, which gives according to equation (5)  $\rho = 1.4$ , which is in rough agreement with the results shown in fig. 8. We may, therefore, conclude that in the above experiments the fibred strips broke at an angle which lay much more nearly symmetrically to the maximum elongation than to the maximum tension.

Describing the angle of break with respect to the direction of maximum elongation, the results shown in Table II are obtained.

Table II.—The Angle of Break Measured from the Direction of Maximum Elongation.

$\theta$ .	Angle of break.
°	°
0	60
15	65
30	73
45	75

Of the two results obtained at high temperatures, the first, that the angle of break depends on grain orientation, seems fairly well established. The second, that the angle of break lies symmetrically to the direction of maximum elongation, rests on rather scanty evidence, and the result is here reported as an observed peculiarity of the actual observations which may or may not prove capable of generalization.

#### *A Simple Test for the Degree of Fibring.*

Since, as is shown in Table II, the angle of break depends on the orientation of the fibre structure, it might be supposed that it likewise depends on the degree of fibring. This was experimentally verified. When sheets are rolled in stacks, some are occasionally found less well fibred than others. Two sheets from the same stack were chosen, of which one showed much more pronounced fibring than the other. Strips cut at  $45^\circ$  to the rolling direction in these two

sheets broke at very nearly the same angle, but of the strips cut parallel to a rolling direction, the well-fibred ones broke at an angle more obtuse by  $15^{\circ}$  than the strips having a random grain orientation.

In conclusion, I should like to express my indebtedness to Mr. A. A. Frey, Dr. V. Hicks, and Mr. P. G. McVetty, of the Westinghouse Research Laboratories, for their valuable assistance in carrying out various parts of the experimental work.

*Summary.*

Experiments are reported on specially prepared iron-silicon sheets which are characterized by a fibre structure, as revealed by X-rays, in which diagonal axes of the crystal grains are disposed symmetrically about two mutually perpendicular directions in the plane of the sheet. Strips cut from such sheets in various directions were broken by tension at the temperature of liquid air and a few degrees above room temperature. At low temperatures the material is brittle, and there is a tendency for fracture to take place along (100) planes only. At the higher temperatures fracture takes place along a line making a certain angle with the direction along which tension is applied. This angle depends on the orientation of the fibre-structure within the strip. It is shown that the angle of break may be used to estimate the degree of fibring in a rolled sheet.

---

### *The Beilby Layer.\**

By G. I. FINCH, A. G. QUARRELL, and J. S. ROEBUCK, Imperial College of Science and Technology.

(Communicated by W. A. Bone, F.R.S.—Received February 21, 1934.)

[PLATE 18.]

#### *Introduction.*

In the course of an experimental study of the relationship between the photoelectric properties and structure of certain surfaces, crystalline zinc films of normal structure were required. In order to avoid pseudomorphic distortion† of the zinc lattice, the films were deposited *in vacuo* on a polished copper substrate. To our surprise we were unable to obtain a zinc film exhibiting any organized structure, although such films, when deposited under identical conditions, but on glass or quartz, otherwise invariably yielded clear and well-defined electron diffraction patterns. Accordingly, the experiments described below were carried out with the object of inquiring into the nature of this phenomenon, and it seems to us that the results confer an objective reality upon the Beilby layer, which raises its existence from the realm of hypothesis to that of established fact.

#### *Experimental.*

The experiments consisted in observing the development of an electron diffraction pattern during condensation of a metal vapour on a substrate, and the pattern changes occurring subsequent to completed deposition.

A tungsten-filament metal evaporating assembly, similar to that previously described‡, was introduced into the diffraction camera§ through a 5 cm. diameter port situated opposite to the specimen carrier. The filament was charged with the metal to be vaporized either by immersion in, and withdrawal from, the melt or by wrapping round it the wire or foil. Prior to deposition the specimen carrier was adjusted to bring the receiver into the reflection

\* Sir George Beilby, "Aggregation and Flow of Solids," Macmillan, London (1921), according to whom polish is the result of surface flow whereby a vitreous, non-crystalline layer, now generally known as the "Beilby Layer," is formed.

† Finch and Quarrell, 'Proc. Roy. Soc.,' A, vol. 141, p. 398 (1933).

‡ Finch and Quarrell, 'Proc. Phys. Soc.,' vol. 46, p. 148 (1934).

position in the electron beam. The filament, placed between about 3 and 4 cm. from the receiver, was then heated for a fraction of a second to a temperature sufficient to ensure the speedy evaporation of a portion of the metal with which it was charged. Part of the vapour condensed on the receiver and, as this meanwhile was being swept by the electron beam, structural changes in the resulting film could be closely followed from its initial stages of formation onwards by means of the diffraction pattern visible on the fluorescent screen. In some experiments a metal sheet was interposed between the filament and the receiver to act as a barrier to radiation heating of the receiver surface, which thus remained at room temperature throughout each experiment. It was subsequently found, however, that the results were unaffected by the presence or otherwise of the screen. Indeed, in view of the short duration of "flashing" and the massive nature of the majority of the receivers employed (2.5 cm. diameter discs, about 2 mm. thick) it is unlikely that, in the absence of the screen, deposition of metal vapour was accompanied by anything more than a rise of a few degrees in the temperature of the receiver surface.

The metals evaporated *in vacuo* were zinc, lead, silver, and tin; they were deposited on surfaces of copper, iron, zinc, lead, or gold. The substrate was either highly polished or consisted of a previously polished surface which had been suitably etched. The first stage in polishing was carried out by high-speed buffing with beeswax and rouge. The final polishing was effected with wet, light-grade magnesia on chamois leather, followed by gentle buffing on chamois leather of the softest grade obtainable. A properly polished surface did not yield any regular diffraction pattern but merely gave rise to random electronic scattering. On the other hand, an unsuitably polished surface yielded a diffraction pattern consisting of two more or less indistinct broad rings with much background scattering. The diffraction pattern characteristic of the metal could be restored by suitable etching of the polished surface. Immediately after polishing, or washing with distilled water to remove the etching reagent, the receivers were placed in the diffraction camera which was then evacuated. In all experiments the cathode chamber leak system was supplied with hydrogen. The electron beam was filtered clear of retrograde rays by magnetic biasing in the manner previously described.\*

*The Results.*—In the first experiment zinc was deposited on polished copper. A brilliant well-defined diffraction pattern characteristic of the structure of normal zinc, fig. 1, Plate 18, was obtained immediately on "flashing." Thereupon the pattern was seen to fade, though without any noticeable loss of ring

\* Finch and Quarrell, 'Proc. Roy. Soc.,' A, vol. 141, p. 398 (1933).

definition, until after about 10 seconds no vestige of the pattern remained. The experiment was continued, so that in all 12 successive deposits of zinc were "flashed" over. With each layer, except the last, the initial clear and well-defined diffraction pattern vanished at a rate decreasing with each successive deposit. For example, the diffraction pattern yielded by the last layer but one was still faintly visible 4 minutes after deposition, but had completely vanished after 5 minutes. After an initial period of weakening the pattern from the 12th deposit remained unchanged, even after 4 hours.

In a second experiment, in which zinc was again deposited on polished copper, the pattern due to the first layer disappeared completely within 3 seconds, during which interval the two patterns, figs. 2, *a* and *b*, Plate 18, were recorded. These, together with fig. 1, show clearly how fading of the diffraction patterns occurred without noticeable broadening of the rings.

Further experiments in which zinc was deposited on polished copper gave results which showed that the time of fading of the diffraction pattern yielded by a freshly formed deposit depended upon the amount of metal "flashed" over and on the manner in which polishing had been carried out.

In another experiment the receiver consisted of a copper disc, one side polished, the other etched. The plane of the disc was in line with that of the metal-evaporating filament. The electron beam grazed the polished surface. Zinc was "flashed" over, whereupon a well-defined normal zinc pattern appeared on the screen, to disappear completely within 20 seconds. The etched side of the receiver then was rotated into position in the beam, and a pattern was obtained which showed neither loss in brilliancy nor other change in appearance during  $1\frac{1}{2}$  hours and possibly longer. On removal of the receiver from the camera, it was found that the layer of zinc deposited on the etched copper surface could be easily removed by gentle rubbing with filter paper. The deposit on the polished substrate, however, not only adhered firmly but also had a brassy colour which could only be removed by abrasion.

The further experiments and results are given in Table I.

In general, a metal layer deposited on a polished substrate adhered far more firmly thereto than on the etched, i.e., crystalline surface. The one exception to this rule was furnished by tin deposited on polished zinc, but in this case the first layer also yielded a permanent pattern. It has previously been shown,\* however, that an oxide film is readily formed on a zinc surface on exposure to air; thus it is almost certain that the polished zinc substrate

\* Finch and Quarrell, 'Proc. Roy. Soc.,' A, vol. 141, p. 398 (1933).



employed in these experiments was in effect covered by a thin oxide film formed between final polishing and transfer to the camera.

*Discussion.*

The effect of polishing of metallic surfaces upon the corresponding electron diffraction patterns has previously been examined by several workers. Kirchner\*

Table I.

Experiment No.	Metal deposited.	Substrate.	Results.
IVa	Zn	Polished mild steel	First layer pattern disappeared within 3 minutes ; second layer pattern weakened to a limit in about 5 minutes ; no further change in ensuing 1½ hours.
IVb	Zn	Etched mild steel	No diminution in zinc pattern intensity during 2 hours.
V	Zn	Polished lead	First layer pattern disappeared within 2½ minutes.
VIa	Zn	Polished gold	First layer pattern disappeared within 2 hours.
VIIb	Zn	Etched gold	First layer pattern unchanged after 15 hours.
VIIa	Sn	Polished copper	First layer pattern disappeared within 5 minutes ; second layer yielded sharp rings which became fainter but were still faintly visible after ½ hour.
VIIb	Sn	Etched copper ....	Pattern unchanged after 12 hours.
VIIIa	Sn	Polished gold	First layer gave a brilliant pattern which disappeared after ½ hour.
VIIIb	Sn	Etched gold . . . . .	Pattern unchanged after 12 hours.
IX	Sn	Polish mild steel	First layer pattern faded in 25 minutes to a much weaker but permanent pattern.
X	Sn	Polished lead .....	Seven successive layers deposited, but no rings visible at any time.
XI	Sn	Polished zinc	Clear first layer pattern obtained which showed no change after 1 hour.
XII	Pb	Polished copper	Faint first layer pattern disappeared rapidly.
XIIIa	Pb	Polished gold	Good first layer pattern disappeared within 10 minutes.
XIIIb	Pb	Etched gold ..	Pattern unchanged after 3 hours.
XIVa	Ag	Polished copper	First layer rings disappeared very slowly.
XIVb	Ag	Etched copper	Pattern unchanged after 2 hours.

considered that the broadening of the rings with increasing degree of polish was due, in the main, to a levelling of the surface, and could not properly be ascribed to crystal attrition. On the other hand, French,† as the result of a detailed and systematic experimental study of polycrystalline and single crystal metallic surfaces, was led to the view that polishing rendered such

\* 'Nature,' vol. 129, p. 545 (1932).

† 'Proc. Roy. Soc.,' A, vol. 140, p. 637 (1933).

surfaces practically amorphous. This conclusion was challenged by Germer\* and by Kirchner and Raether†; but Raether‡ later withdrew his criticism.

It has been shown above that, though well-polished surfaces of the metals examined by us failed to yield ordered diffraction patterns, suitable etching sufficed to restore such surfaces to a condition enabling patterns characteristic of their normal crystalline structures to be obtained. Further, we have found that suitable crystalline metal films deposited on these surfaces lost their structure on polished, but not on crystalline substrates; and that the rate of disappearance of structure was largely determined by the method of polishing. Finally, in some experiments, for example, in the case of zinc deposited on polished mild steel, saturation of the substrate polish layer with the deposited crystalline metal has been demonstrated.

It is our view that these facts support the late Sir George Beilby's conception of the effects of polishing, according to which polishing destroys the crystalline structure of a surface and causes a molecular flow, whereby the surface becomes covered with an amorphous layer which may properly be likened to a highly viscous liquid.

The direct nature of the experiments outlined above may well be illustrated by the following analogy: when the diffraction pattern of the crystalline film freshly deposited on a polished surface is seen to fade it is as if we were observing through a powerful microscope the solution of crystals supernatant on a solvent. The fact that the diffraction rings were seen to fade without broadening proves that erosion of the crystals was confined to faces in contact with the substrate.

We wish to thank the Government Grant Committee of the Royal Society for a grant for apparatus, the Cambridge Instrument Company for the electron diffraction camera and the Department of Scientific and Industrial Research for a grant which has enabled one of us (A. G. Q.) to devote his whole time to this research.

Metal vapours have been condensed on polished and etched metal surfaces under observation in an electron diffraction camera.

It has been found that (i) suitably polished surfaces gave rise solely to random electronic scattering; (ii) etched surfaces exhibited a crystalline structure;

\* 'Phys. Rev.,' vol. 43, p. 724 (1933).

† 'Phys. Z.,' vol. 33, p. 510 (1932).

‡ 'Z. Physik,' vol. 86, p. 82 (1933).



FIG. 1. Electron diffraction pattern from freshly deposited zinc film.  
FIG. 2*a*. 1 second after deposition of zinc on polished copper.  
FIG. 2*b*. 3 seconds after deposition of zinc on polished copper



(iii) the diffraction pattern yielded by a crystalline deposit disappeared more or less rapidly when the substrate was polished, but was permanent on an etched surface; (iv) fading of a diffraction pattern occurred without noticeable broadening of the rings; (v) when several layers were formed on a polished surface, the corresponding diffraction patterns vanished at rates decreasing with each successive deposit, and in some experiments the pattern due to the final deposit only weakened to a limit at which it remained unchanged.

It is concluded that these facts afford a direct experimental proof of the existence of the Beilby layer and of its formation by polishing.

### *Investigations on the Spectrum of Selenium. III.—Extension of Se III.*

By K. R. RAO, D.Sc., and S. G. K. MURTI, M.A., Research Scholar, Andhra University, Waltair, India.

(Communicated by A. Fowler, F.R.S.—Received February 27, 1934.)

In the second of these contributions dealing with investigations on the successive spectra of selenium,\* evidence was obtained that the spectrum of Se III consists of singlets and triplets, and nearly all the terms corresponding to the deepest  $4p$  and the higher  $5s$ ,  $4d$ ,  $5p$  and  $sp^3$  states were discovered. The present work is an extension of the identification of the predicted terms, particularly those arising from the still higher states  $6s$  and  $5d$ ; it also gives the results of a further investigation of the spectrum in the infra-red region.

It was pointed out in Part II that many of the lines due to the transition  $4d \rightarrow 5p$  must be situated in the infra-red above  $\lambda 6700$ , and that their wavelengths could be exactly calculated from the known values of the terms. The spectrum in this region has accordingly been investigated with a Hilger Constant Deviation Spectrograph. The source employed was the usual condensed discharge through a capillary tube excited by a  $\frac{1}{2}$  K.W. transformer. To photograph this region, krypto- and neo-cyanine plates were used, after hypersensitizing them in a bath of ammonia. Exposures extended over a

\* K. R. Rao and Badami, 'Proc. Roy. Soc.,' A, vol. 131, p. 154 (1931); Badami and K. R. Rao, *ibid.*, vol. 140, p. 387 (1933). These will be referred to subsequently as Part I and Part II.

period of 8 to 10 hours. The lines of the iron arc served as standards for the measurement of the plates.

Table I gives the wave-lengths of the lines observed and of those predicted in Part II; the agreement between these is a strong support to the classification proposed previously.

Besides the above combinations in the infra-red, the exact positions of a few more multiplets could be calculated from the scheme of terms given in Part II. These are due to the transition  $sp^3 \rightarrow 5p$ . An examination of the

Table I.

$\lambda$ calculated.	$\lambda$ observed.	$\lambda$ calculated.	$\lambda$ observed.
9563.83		7761.12	
9516.39		7349.40	7352
9289.80		7322.22	7323
9243.99		7243.49	7246
9175.50		7147.63	
8688.50	8689	7117.91	7119
8313.45	8315	7084.62	
8001	8004	7068.54	
7952.19	7955	6810.90	

discharge tube plates in the Schumann region partially revealed the existence of the groups  $sp^3$ ,  $^3P$ ,  $^3D \rightarrow 5p$ ,  $^3P$ ,  $^3D$ . These are shown in the multiplet tables given below.

While attempting to discover the terms  $^1^3D$ ,  $^1^3P$ ,  $^1^3S$  of the  $5p$  state, a difficulty was met with (*cf.* Part II, p. 391) regarding the assignment of the level  $\nu = 118563.6$ , which was finally designated as  $5p^3P_1$ , in spite of the consequent abnormal intensity of the intercombination line  $5s^1P_1 - 5p^3P_1$ . This conclusion was arrived at after correlating the iso-electronic spectra of As II\* and Se III. It was mentioned also that the level  $5p^1P_1$  was undiscovered. A further search has now revealed a new level  $\nu = 124162.9$  which could, possibly, only be denominated as  $5p^1P_1$ . It will be seen that there is again an abnormally large intensity of the inter-combination lines  $5s^3P_1$ ,  $^3P_1 - 5p^1P_1$ , as compared with that of  $5s^1P_1 - 5p^1P_1$ . Effective comparison of these relative intensities cannot be made with As II, as the corresponding singlet line of As II is situated in an inaccessible region in the infra-red. But it may be observed that there is a close analogy between As II and Se III regarding the relative magnitude of the levels of the  $5p$  state; the  $5p^1P_1$  in

\* A. S. Rao, 'Proc. Phys. Soc., Lond.,' vol. 44, p. 343 (1932); also 'Ind. J. Phys.,' vol. 7, p. 561 (1933).

both these spectra being the deepest of all the triplets and singlets due to this configuration.

It might be suggested that the interchange of the levels  $5p\ ^1P_1$  and  $5p\ ^3P_1$  would remove the irregularities in the intensities of the lines referred to above, but there would then result entirely anomalous  $^3P$  term intervals. The writers consider that the latter anomaly is more improbable.

The classifications discussed above, together with the combinations between the terms  $5p\ 6s$ ,  $5p\ 5d$  and  $4p\ 6s$ ,  $5d$ , are set forth in the usual multiplet form in Table III. The levels  $5d\ ^3P_0$  and  $5d\ ^1P_1$  could not be satisfactorily located, as also  $4p\ ^1S_0$  and  $5p\ ^1S_0$ . A very exhaustive search has been made to identify the level  $4p\ ^1S_0$  as it is one of the ground terms of the spectrum, but with no definite result. The suggestion that

$$4p\ ^1S_0 - 5s\ ^1P_1 = 104213$$

$$4p\ ^1S_0 - 5s\ ^3P_1 = 98697$$

may prove to have no significance, for the first of these lines must then be assumed to be coincident with the strong combination  $4p\ ^3P_1 - 4p\ ^3D_1$ , of Se IV.

For convenience of reference, a brief statement of the predicted terms of Se III is given in Table II.

Table II.—Predicted Terms of Se III.

Electron configuration.	Terms.
$4p$	$^3P\ ^1D\ ^1S$
$5s$	$^3P\ ^1P$
$4d$	$^3F\ ^3D\ ^3P\ ^1F\ ^1D\ ^1P$
$5p$	$^3D\ ^3P\ ^3S\ ^1D\ ^1P\ ^1S$
$sp^3$	$^3S\ ^3S\ ^3D\ ^3P\ ^1D\ ^1P$

It must be stated that although there can be no doubt as to the reality of the various levels of the  $6s$  and  $5d$  states that have been identified, the designations of these levels suggested here cannot be claimed to be absolutely certain, as they can be based only on such general considerations as those of the relative intensities of the lines forming a multiplet and the relative magnitudes of the intervals. However, the detection of many of the calculated lines due to the combinations  $4p \rightarrow 6s$  and  $\rightarrow 5d$  exactly in the predicted positions, in the region  $\lambda\ 500 - \lambda\ 600$ , with the appropriate intensities, lends strong support to





$^1P_2 = 78078.6$	—	—	38870.9	35675.2 (8)	
$^3D_1 = 84082.8$ — 821.0	37320.5 (2)		—	29670.8 (0)	40080.1 (6)
$^3D_2 = 84903.8$ — 3894.8	36499.4 (2)	33327.9 (2)	32146.8 (2)	28850.0 (1)	39259.7 (2)
$^3D_3 = 81009.0$	—	37222.1 (6)	36040.9 (7)	—	
$^1D_2 = 81619.6$	39783.5 (2)	36811.6 (0)	35431.5 (4)	34002.1 (0)	42543.8 (4)
$^3P_0 = —$	—	—	—	—	—
$^3P_1 = 79973.5$ — 223.5	—	—	37076.2 (3)	35647.9 (3)	—
$^3P_2 = 80197.0$	—	38034.1 (2)	36851.9 (4)	35424.6 (5)	—

Table III—(continued).

$4p$	$^3P_0$ 274924	$^3P_1$ 1741 273183	$^3P_2$ 2196 270987	$^1D_2$ 261892
$6s\ ^3P_0 = 87755.3$ $^{257.8}$		185422 (0)		
$^3P_1 = 87497.5$ $^{4096.5}$	—	185691 (2)	183486 (1)	174386 (0)
$^3P_2 = 83401.0$		—	187593 (1)	—
$^1P_1 = 82762.2$	—	190429 (0)	—	179128 (1)
$5d\ ^3F_3 = 86494.8$ $^{1218.7}$	—	—	—	—
$^3F_3 = 85276.1$ $^{1945.7}$			—	176604 (00)
$^3F_4 = 83330.4$				
$^1F_3 = 78078.6$			—	183813 (2)
$^3D_1 = 84082.8$ $^{-821.0}$	190836 (2)	—	—	—
$^3D_2 = 84903.8$ $^{3894.8}$		188274 (3)	186088 (0)	176979 (0)
$^3D_3 = 81009.0$			189970 (4)	—
$^1D_2 = 81619.6$		191553 (1)	—	180268 (1)
$^3P_0 =$ —				
$^3P_1 = 79973.5$ $^{223.5}$	—	193211 (2)	191011 (0)	—
$^3P_2 = 80197.0$		192987 (0)	190785 (3)	181686 (0)

$$5s\ ^3P_0 - 5p\ ^1P_1 = 24483.6 (7)$$

$$^3P_1 - ^1P_1 = 23979.9 (10)$$

$$^1P_1 - ^1P_1 = 19105.2 (5)$$

$$4d\ ^3F_3 - ^1P_1 = 26709.0 (8)$$

the correctness of the assignments. Further confirmation is afforded, particularly, at least for the  $6s$  terms, by the following discussion of the various groups of terms.

It is seen that, in spite of the interval ratios of the  $ms\ ^3P$  term being large, the total magnitudes of the intervals exhibit a slow and gradual progression towards the limiting value  $4376\text{ cm.}^{-1}$  of  $4p\ (^3P_2 - ^3P_1)$  of Se IV.

$5s\ ^3P_0$			$6s\ ^3P_0$		
	504.5	] 4114.1		257.8	] 4354.3
$^3P_1$	3609.6		$^3P_1$	4096.5	
$^3P_2$	1264.9		$^3P_2$	638.8	
$^1P_1$			$^1P_1$		

There is further a diminution of  $^3P_0$ — $^3P_1$  and  $^3P_2$ — $^1P_1$  and an increase of  $^3P_1$ — $^3P_2$  as we proceed to the higher members of the series; this leads to the conclusion that  $ms\ ^3P_0$  and  $^3P_1$  should converge to  $4p\ ^3P_{1\frac{1}{2}}$ , while  $ms\ ^3P_2$  and  $^1P_1$  converge to  $4p\ ^3P_{1\frac{1}{2}}$  of Se IV. This behaviour is in keeping with the theoretical conclusions of Hund and Goudsmit. It was, however, also predicted that, in spectra of the type under consideration, the separation of the  $s$  levels with  $j = 2$  and  $j = 0$  is exactly constant through the whole series of levels and equal to the “ $p$ ” doublet of the ion, independently of the type of coupling.\* The slightly smaller value of the triplet separations obtained in Se III, as well as in many of the other similar spectra, must therefore be ascribed to the screening action of the second valence electron, which has the effect of reducing the interval.

The manner in which the series terms coming from the “ $md$ ” electrons approach their limits is difficult to determine, since these terms present no very obvious regularity of any kind, as is evident from the following table of these intervals in all the similar spectra, isoelectronic with C I, Si I, and Ge I, which have hitherto been investigated.

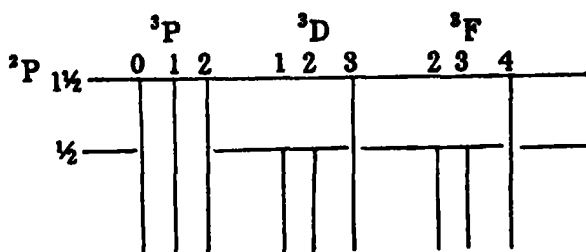


FIG. 1.

From theoretical considerations Hund has concluded that the limits of the terms arising from the addition of an “ $md$ ” electron can be represented by the scheme shown in fig. 1.

According to this representation the intervals of the  $^3P$  term and also the differences  $^3D_1$ — $^3D_2$  and  $^3F_2$ — $^3F_3$  should progressively diminish with increasing

\* ‘Phys. Rev.’ vol. 31, p. 946 (1923).

Table IV.—“ $md$ ” Term Intervals in  $s^2p, d$  Electron Configuration.

Term.	Intervals.			Intervals.			Intervals.		
	C I.	N II.	O III.	Si I.	P II.	S III.	Ge I.	As II.	Se III.
$md^4F_1$	15.7?	—	195.79	—	—	—	254.1	722	1257.0
$^4F_3$	34.5?	—	178.16	—	—	—	—	—	2099.7
$^4F_4$									
$^3D_1$	5.0	24.1	49.24	19.9	20.3	140.89	-80.4	483	-1230
$^3D_2$	9.2	30.3	73.72	52.0	34.7	53.39	262.2	842	2604.9
$^3D_3$									
$^3P_0$	-4.3	-28.1	-61.45	unresolved	-10.8	20.2	-275.2	80	442.4
$^3P_1$	-9.5	-51.8	-114.00		-48.4	7.8	-267.2	-475	-52.1
$^3P_2$									
$(m+1)d^4F_1$	10.5	—	—	—	—	491.86	—	—	1218.7
$^4F_3$	—	—	—	—	—	489.82	—	—	1945.7
$^4F_4$									
$^3D_1$	6.4	26.5	—	55.5	—	132.74	-97.3	—	-821.0
$^3D_2$	10.5	35.6	—	87.0	—	239.36	313.2	—	3894.8
$^3D_3$									
$^3P_0$	—	-25.5	—	—	—	152.7	-265.4	—	—
$^3P_1$	-10.1	-46.1	—	—	—	771.5	-252.2	—	-223.5
$^3P_2$									

values of  $m$ , while  $^3D_2$ — $^3D_3$  and  $^3F_3$ — $^3F_4$  should increase up to their limiting values. No long series have been determined in any spectrum to justify an attempt to draw definite conclusions from the available experimental data. However, some general features may deserve notice. Two terms of the series  $md^3D$  have been found in many of these spectra. It will be observed that in all of them the total term separation  $^3D_1$ — $^3D_3$  increases when we pass to the corresponding higher member. This would mean that the two components  $^3D_1$  and  $^3D_2$  tend separately towards the two limiting levels  $^3P_1$  and  $^3P_2$ , respec-

tively, which is in keeping with the above predictions. At the same time it is difficult to conclude about the limit of the series  $md^3D_3$ . Perhaps the increasing values of  $^3D_3$ — $^3D_3$  and the partial inversion of the D term in certain cases indicate the correctness of Hund's conclusions. Similar conclusions cannot be arrived at for the  $md^3P$  term. With regard to the  $^3F$  term, however, Hund predicts the limits

$$\begin{aligned} ^3F_3, ^3F_2 &\rightarrow ^3P_1 \\ ^3F_4 &\rightarrow ^3P_{11} \end{aligned}$$

whereas, at least for Se III (in which alone two  $md^3F$  terms have been identified) the assignments made in the present work, if correct, lead to the suggestion that all the components of the  $^3F$  term appear to converge to the same limit, presumably  $^3P_{11}$ . This calls to mind a somewhat similar departure in which all the components of  $d^4F$  and  $d^4D$  terms have been found tending towards the same level  $^3P_2$  of the next higher ion,\* while theoretical predictions† represented in fig. 2 suggest different limits to the component levels.

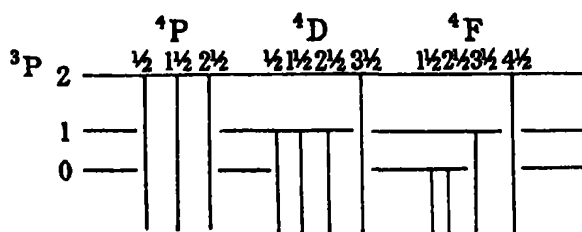


FIG. 2.

It might be instructive to examine the limits for terms arising from the addition of an "mp" electron for which Hund's scheme of limits is shown in fig. 3.

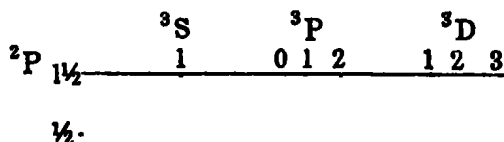


FIG. 3.

Contrary to this scheme, the intervals of  $4p^3P$  and  $5p^3P$  in Se III given in Table V clearly show that all the components of these terms converge to a common limit.

\* Kees and de Bruin, 'Bur. Stand. J. Res.,' vol. 2, p. 1133 (1929).

† Hund, "Linienspektren und Periodisches System der Elemente," p. 198 and fig. 34.

Table V.— $mp\ ^3P$  Intervals in Se III.

$4p\ ^3P_2$	] 2196 ]	3937	$5p\ ^3P_2$	] 1513.8 ]	3091.2
$^3P_1$			$^3P_1$		
$^3P_0$			$^3P_0$		

An exactly similar behaviour is noticeable in the variation of  $mp\ ^3P$  differences in all the spectra of this type which have so far been analysed. The tendency of  $mp\ ^3D$  terms cannot be judged, as only one of these terms has been identified in all these spectra.

Turning to other features which the above comparisons would reveal, it might be worth noting that, as would generally be expected on account of deviations from the usual (LS) type of coupling, an increasing nuclear charge always distinctly favours deviations from the normal characteristics of the spectrum. Again, of a set of terms arising from a given configuration, the higher lying terms appear to exhibit greater anomalies. For instance (*cf.* Table IV), the partial inversion of the term and abnormal ratios are more marked in  $^3D$  and  $^3P$  than in  $^3F$  of the group of  $md$  terms.

#### *Term Values.*

It is possible to make a new and independent estimate of the relative term values in Se III with the aid of the newly discovered groups. But, as more than two members of any series have not been identified, no greater accuracy can be expected than was previously obtained. The term values given in Table VI are therefore based on the value of  $4p\ ^3P_0 = 274924$  adopted in Part II, and this table contains, with the exception of  $4d\ ^3F$ ,\* only those terms which are newly evaluated.

#### *Catalogue of Lines of Se III.*

The observations included in the present catalogue (Table VII) extend from the red to  $\lambda\ 400$  in the extreme ultra-violet, and contains all the lines of selenium in that region, which have appeared to be attributable to Se III, with the exception of those already classified and listed in Part II, Table V; these latter are omitted here to avoid duplication.

As experimental evidence alone, for the distinction of lines, is not very conclusive, it is possible that some of the unclassified lines in Table VII may

\* Owing to an error in numerical calculation the values of  $4d\ ^3F_{2,3,4}$  were wrongly given in Part II as 135863.7, 134405.8, and 132306.1 respectively. The present corrected values should be adopted.

Table VI.—New Terms in Se III.

Term.	Term value.	Term.	Term value.
$sp^1D_2$	162359	$5d^1F_3$	86494.8
$4d^1F_2$	150872.1	$^1F_3$	85276.1
$^1F_2$	149614.2	$^1F_4$	83330.4
$^1F_4$	147614.5	$^1D_1$	81619.6
$5p^1P_1$	124162.9	$^3D_1$	84082.8
$6s^1P_1$	82762.2	$^3D_1$	84903.8
$^3P_0$	87755.3	$^3D_2$	81009.0
$^3P_1$	87497.5	$^3P_0$	—
$^3P_2$	83401.0	$^3P_1$	79973.5
$5d^1F_2$	78078.6	$^3P_2$	80197.0

Table VII.—Catalogue and Classification of Se III Lines.

$\lambda$ (int.) Å.	$\nu$ (vac.)	Classification.	$\lambda$ (int.) Å.	$\nu$ (vac.)	Classification.
6545.24 (2)	15274.1	$5s^1P_1-5p^1P_1$	3904.00 (5)	25607.5	$5p^1D_2-6s^1P_1$
6448.89 (1)	15502.3		3901.55 (10)	25623.6	
6417.19 (3)	15578.9		3883.39 (1)	25743.4	
6291.12 (0)	15891.0		3858.09 (5)	25912.2	
6239.65 (2)	16022.1				
6125.76 (2)	16320.0		3857.29 (9)	25917.6	
6116.28 (1)	16345.3		3853.30 (8)	25944.5	
6106.06 (1)	16372.7		3827.81 (1)	26117.2	
5959.75 (0)	16774.6		3812.16 (7)	26224.4	
5825.10 (1)	17162.3		3807.54 (1)	26256.3	
5232.73 (5)	19105.2		3744.01 (3)	26701.8	
5150.09 (1)	19411.7		3742.99 (8)	26709.0	
4651.51 (0)	21492.4		3667.58 (1)	27258.2	
4554.00 (3)	21952.6		3654.91 (3)	27352.7	
4523.52 (8)	22100.5		3654.69 (1)	28123.8	$^3S_1-6s^1P_1$
4420.81 (0)	22614.0	$^3P_1-^1P_1$	3648.96 (1)	28169.2	
4251.76 (2)	23513.1		3648.59 (0)	28172.2	
4234.45 (1)	23609.2		3516.94 (2)	28425.7	$^1D_2-5d^1D_2$
4186.29 (5)	23880.8		3493.89 (00)	28613.2	
4184.90 (6)	23888.7		3485.91 (2)	28678.7	
4169.10 (10)	23979.9		3465.22 (1)	28850.0	
4165.67 (1)	23999.0		3427.87 (1)	29164.3	
4153.93 (4)	24066.8		3382.91 (1)	29551.9	
4149.05 (4)	24085.1		3376.26 (4)	29610.1	
4145.31 (4)	24116.9		3369.35 (0)	29670.8	
4113.53 (3)	24309.1		3362.83 (1)	29728.3	
4097.92 (6)	24395.8		3342.40 (1)	29910.1	
4083.21 (7)	24483.6		3318.07 (1)	30129.4	
4059.83 (5)	24624.6		3293.67 (2)	30352.6	$5p^1D_2-6s^1P_1$
4014.00 (8)	24905.8		3292.61 (0)	30362.3	
4008.27 (10)	24941.4		3271.90 (1)	30554.3	
3993.73 (8)	25032.3		3245.04 (2)	30807.4	
3981.34 (1)	25110.1		3244.36 (2)	30813.9	
3963.93 (1)	25220.4		3242.79 (5)	30828.8	
3957.43 (0)	25261.8		3236.55 (1)	30888.2	
3937.40 (0)	25390.3		3225.82 (5)	30991.0	
3935.90 (10)	25401.3		3218.05 (2)	31065.8	$^1D_2-^1P_1$
3931.57 (10)	25428.0		3178.20 (6)	31455.3	
3904.86 (8)	25601.9		3117.60 (2)	32066.7	

Table VII—(continued).

$\lambda$ (int.) I.A.	$\nu$ (vac.).	Classification.	$\lambda$ (int.) I.A.	$\nu$ (vac.).	Classification.
3111.05 (4)	32134.2	$^1D_2-5d\ ^1D_2$	2565.25 (0)	38970.9	$^3P_1-5d\ ^1F_3$
3109.93 (2)	32145.8	$^3P_1-^1D_2$	2546.38 (2)	39259.7	$^1P_1-^1D_2$
3102.76 (5)	32220.0	$^3S_1-6s\ ^3P_1$	2544.56 (0)	39287.7	
3063.75 (4)	32630.3		2513.37 (1)	39775.2	
3062.55 (3)	32643.1	$^3P_0-^3P_1$	2512.85 (2)	39783.5	$^3D_2-^1D_2$
3051.04 (1)	32766.2		2494.25 (6)	40080.1	$^1P_1-^1D_2$
3042.45 (0)	32858.7	$^3S_1-^1P_1$	2493.45 (3)	40093.0	$^3D_1-^1D_2$
3033.58 (1)	32954.8	$^3D_2-5d\ ^1F_3$	2464.61 (0)	40562.1	
3031.48 (0)	32977.6		2459.54 (6)	40645.7	$sp\ ^1D_2-5p\ ^3D_1$
3020.27 (1)	33100.0		2452.52 (0)	40762.0	$5p\ ^1P_1-6s\ ^3P_1$
2999.62 (2)	33327.9	$^1D_2-^1D_2$	2440.94 (3)	40955.4	$sp\ ^1D_2-5p\ ^3D_1$
2987.45 (2)	33463.6		2435.52 (1)	41046.5	
2979.09 (2)	33557.5	$^1D_2-^3P_1$	2420.27 (5)	41305.1	
2970.95 (6)	33649.5	$^3P_1-6s\ ^3P_1$	2414.77 (2)	41399.2	$5p\ ^1P_1-6s\ ^1P_1$
2969.96 (4)	33660.7	$^3P_1-5d\ ^1D_2$	2400.13 (3)	41651.7	
2948.46 (6)	33906.1	$^3D_2-6s\ ^3P_1$	2391.93 (0)	41794.5	
2947.84 (2)	33913.2		2390.06 (4)	41827.2	
2944.02 (4)	33957.3	$^3D_1-^3P_0$	2371.06 (6)	42162.3	
2940.14 (0)	34002.1	$^3S_1-5d\ ^1D_2$	2353.75 (2)	42472.4	
2937.11 (00)	34037.1		2349.80 (3)	42543.8	$^1P_1-5d\ ^1D_2$
2921.83 (2)	34215.1	$^3D_1-6s\ ^3P_1$	2307.62 (0)	43321.3	
2899.27 (1)	34481.4	$^3P_1-5d\ ^1D_2$	2282.62 (4)	43795.8	$sp\ ^1D_2-5p\ ^3P_1$
2870.17 (6)	34830.9	$5p\ ^1D_2-6s\ ^3P_1$	2263.12 (0)	44173.0	$^3P_1-^1P_1$
2864.43 (6)	34900.7	$^3D_2-5d\ ^1F_3$	2259.28 (2)	44248.1	$^3P_1-^1P_1$
2863.84 (6)	34907.9	$^3D_2-^1F_3$	2231.39 (0)	44801.2	
2843.13 (1)	35162.2	$^3P_1-6s\ ^3P_1$	2201.54 (1)	45408.5	
2838.68 (5)	35217.8	$^3D_1-5d\ ^1F_3$			
2822.07 (5)	35424.6	$^3S_1-^3P_1$	$\lambda$ (vac.).		
2821.52 (4)	35431.5	$^3P_1-^1D_2$	2145.18 (1)	46616	
2804.39 (3)	35647.9	$^3S_1-^3P_1$	2141.68 (4)	46692	
2802.24 (8)	35675.2	$^1D_2-^1F_3$	2139.63 (2)	46737	$sp\ ^1D_2-5p\ ^3S_1$
2793.13 (3)	35791.6		2127.48 (2)	47004	$^3P_1-^1D_2$
2792.34 (3)	35801.7	$^3P_1-6s\ ^1P_1$	2111.76 (7)*	47364	
2777.52 (8)	35992.7		2106.32 (4)†	47476	
2773.81 (7)	36040.9	$^3P_1-5d\ ^3D_2$	2101.20 (1)†	47592	
2772.46 (5)	36058.4	$^3P_0-^1D_2$	2057.49 (6)	48603	$^1D_2-^1D_2$
2767.20 (8)	36127.0	$^3D_2-^1F_3$	2009.24 (2)	49770	$^3P_1-^3P_1$
2745.86 (2)	36407.7	$^1P_1-6s\ ^3P_0$	2006.30 (6)	49843	$^3P_1-^3P_1$
2738.96 (2)	36499.4	$^3D_2-5d\ ^3D_2$	2005.05 (2)	49874	$^3P_0-^3P_1$
2730.57 (0)	36611.6	$^3D_2-^1D_2$	1993.02 (2)	50175	$^3P_1-^3D_2$
2726.55 (2)	36665.5	$^1P_1-6s\ ^3P_1$	1950.08 (5)	51280	$^3P_1-^3P_1$
2715.94 (5)	36808.8	$^3D_1-5d\ ^3D_2$	1947.19 (5)	51356	$^3P_1-^3P_1$
2712.73 (4)	36851.9	$^3P_1-^3P_1$	1920.34 (6)*	52074	
2705.96 (4)	36944.5	$^3P_1-^1D_2$	1909.91 (1)*	52359	
2696.35 (3)	37076.2	$^3P_1-^3P_1$	1897.25 (4)†	52708	$^3P_1-^3S_1$
2689.12 (2)	37175.9		1894.44 (4)	52786	$^3P_1-^3S_1$
2685.78 (6)	37222.1	$^3D_2-^3D_2$	1893.19 (5)	52821	$^3P_0-^3S_1$
2678.70 (2)	37320.5	$^3D_2-^1D_2$	1891.16 (6)†	52878	
2674.50 (1)	37379.1	$^3P_0-6s\ ^1P_1$	1887.66 (2)*	52976	
2659.38 (0)	37591.6		1885.30 (2)*	53042	
2656.68 (1)	37629.8	$^3D_1-5d\ ^3D_1$	1844.97 (3)	54202	
2654.92 (1)	37654.7		1829.76 (1)	54652	$^3P_1-^1D_2$
2654.04 (6)	37667.2	$^1P_1-^1F_3$	1844.87 (1)	90795	
2628.44 (2)	38034.1	$^3D_2-^3P_1$	1571.51 (3)	63633	$^3D_2-^3P_1$
2617.33 (6)	38195.5	$sp\ ^1D_2-5p\ ^1P_1$	1535.00 (2)	65147	$^3D_1-^3P_1$
2605.69 (0)	38366.1	$5p\ ^3P_1-5d\ ^3P_1$	1453.98 (2)†	68777	
2590.58 (2)	38589.9	$5p\ ^3P_1-5d\ ^3P_1$	1206.53 (6)	82682	
2587.19 (0)	38640.4	$^3D_2-6s\ ^1P_1$	1100.47 (7)	90870	
2580.97 (1)	38733.5				
2571.30 (5)	38879.2				
2566.60 (2)	38930.4	$^3D_2-^1P_1$			



Table VII—(continued).

$\lambda$ (vac.).	$\nu$ (vac.).	Classification.	$\lambda$ (vac.).	$\nu$ (vac.).	Classification.
1100.36 (8)	99964		640.54 (3)*	156118	
953.92 (2)	104831		631.11 (3)*	158451	
891.48 (8)	112173		630.99 (3)*	158491	
890.52 (0)	112294		630.87 (4)*	158562	
			588.82 (0)	169831	
887.45 (3)	112682		573.44 (0)	174386	$4p\ ^1D_2-6s\ ^3P_1$
882.13 (5)	113362		566.24 (00)	176604	$^1D_2-5d\ ^3F_4$
849.57 (0)	117707		565.04 (0)	176979	$^1D_2-^3D_2$
836.06 (5)	119609		558.26 (1)	179128	$^1D_2-6s\ ^3P_1$
833.14 (5)	120028		554.73 (1)	180268	$^1D_2-5d\ ^1D_2$
828.47 (6)	120702		550.40 (0)	181686	$^1D_2-^3P_2$
818.60 (1)‡	122180		545.00 (1)	183486	$^3P_1-6s\ ^3P_1$
818.10 (1)	122234		544.03 (2)	183813	$^1D_2-5d\ ^3F_4$
816.97 (1)	122404		539.31 (0)	185422	$^3P_1-6s\ ^3P_0$
800.49 (1)	124924		538.53 (2)	185691	$^3P_1-^3P_1$
798.10 (6)*	125298		537.38 (0)	186088	$^3P_1-5d\ ^1D_2$
782.84 (5)*	127740		533.07 (1)	187593	$^3P_1-6s\ ^3P_1$
775.77 (0)	128904		531.14 (3)	188274	$^3P_1-5d\ ^1D_2$
774.41 (0)	129131		530.56 (0)	188480	
764.58 (0)	130791		530.12 (0)	188637	
747.07 (1)	133856		529.41 (1)	188900	
745.70 (1)	134102		526.40 (4)	189970	$^3P_1-^1D_2$
718.71 (0)	139138		525.13 (0)	190429	$^3P_1-6s\ ^3P_1$
692.16 (1)*	144475		524.15 (3)	190785	$^3P_1-5d\ ^3P_1$
690.48 (1)	144827		524.01 (2)	190836	$^3P_0-^3D_1$
688.95 (1)*	145148		523.53 (0)	191011	$^3P_1-^3P_1$
674.47 (5)	148265		522.05 (1)	191553	$^3P_1-^1D_2$
644.79 (4)*	155089		518.17 (0)	192997	$^3P_1-^3P_1$
642.70 (5)*	155594		517.57 (2)	193211	$^3P_1-^3P_1$
640.87 (4)*	156038				

\* Probably lines of Se IV.

† Probably lines of Se II.

‡ Close to an arc line.

§ Close to a line of higher stage.

be due to other stages of ionization of the element, say, Se II, on the one hand, or Se IV on the other. Such of these lines about the assignment of which a doubt really exists are indicated in the table itself and explained in footnotes.

There are groups of rather strong lines, particularly between  $\lambda$  2200 and  $\lambda$  1700 and a few faint lines between  $\lambda$  1000 and  $\lambda$  850, which possibly belong to Se III. These are, however, provisionally assigned to Se II and are therefore omitted from the present list, pending the analysis of Se II, which will be dealt with in a succeeding part.

In conclusion, it is a pleasure to express our gratitude to Professor A. Fowler, F.R.S., for his continued interest in our work and valuable help.

A further study of the spectrum of selenium has revealed about 15 new levels assignable to the doubly-ionized atom and arising chiefly from the  $6s$  and  $5d$  electron configurations; about 120 lines have thus been newly classified. An investigation of the infra-red region of the spectrum has led to the confirmation of the  $5s$  and  $4d$  terms previously identified. The limits to which the component levels tend to converge have been discussed.

*Investigations on the Spectrum of Selenium. IV.—Se I and Se VII.*

By K. R. RAO, D.Sc., and S. G. K. MURTI, M.A., Andhra University, Waltair, India.

(Communicated by A. Fowler, F.R.S.—Received February 27, 1934.)

*Se I.*

In the course of these investigations, which are mainly directed towards a study of the characteristic energy levels of the spark spectra of selenium, several lines have been obtained, which from experimental evidence, could be assigned definitely to the arc spectrum of the element. For the complete excitation of the arc spectrum, recourse must be had to special methods such as the addition of a foreign gas used by Runge and Paschen,\* or an electrodeless discharge in the manner adopted by Gibbs and Ruedy.† The sources in our investigations were chiefly discharges through narrow capillary tubes with varying amounts of inductance and capacity, vacuum sparks and a vacuum arc with selenium tipped graphite or aluminium electrodes. It has been found that many arc lines are excited even in the ordinary uncondensed discharge and in the vacuum arc. However, the arc in an atmosphere of nitrogen in the manner adopted previously for Si, Ge, As‡ has been specially used to isolate, in the fluorite region, the lines belonging to the neutral atom. The data presented in this Part have been based on measurements of plates taken with all these sources.

\* 'Astrophys. J.,' vol. 8, p. 70 (1898).

† 'Phys. Rev.,' vol. 40, p. 204 (1932).

‡ Fowler, 'Proc. Roy. Soc.,' A, vol. 123, p. 422 (1929); also K. R. Rao, *ibid.*, vol. 124, p. 465 (1929); vol. 125, p. 238 (1929).

In the arc spectrum of selenium, Runge and Paschen discovered a sharp and a diffuse series systems of what were then considered to be triplets. Later, McLennan, McLay and McLeod\* brought these quintet groups into consonance with the predictions from Hund's theory, and by studying the arc *in vacuo* in the fluorite region, identified combinations involving the low-lying triplet levels. Recently, Gibbs and Ruedy have detected a few more combinations, in particular the member  $5s\ ^3S-5p\ ^3P$  in the extreme red which has enabled them to establish a connected system of term values. With the aid of the observations made by us it has been possible to assign about 50 lines between  $\lambda\ 3619-\lambda\ 1414$  to the neutral atom of Se, and to detect among these a number of new energy levels of Se I.

The spectrum in question, like that of O I, consists of numerous terms of odd multiplicity, the deepest being an inverted  $^3P$  term. Some of the important terms are  $4p\ ^3P, ^1D, ^1S$ ;  $5s\ ^6S, ^3S$ ;  $4d\ ^5D, ^3D$  and their higher Rydberg members.

In Table I only the multiplets due to the transition into the terms of the deepest  $4p$  state are shown. For the others such as  $5s\ ^6S-mp\ ^5P$  and

Table I.

$4p$	$^3P_2$ 76125	$^3P_1$ 76670	$^3P_0$ 78669	$^1D_2$ 69083	$^1S_0$ 55289
$5s\ ^6S_1 = 30476$		46193.2 (6)	48183.0 (10)		
$^6S_1 = 27661$	48463.2 (10)	49008.1 (15)	50997.8 (20)	41421.2 (8)	27626.2 (2)
$4d\ ^5D_1 = 16979$	59146.4 (5)	59691.5 (5)	61680.4 (2)	52105.0 (8)	38310.6 (1)
$^5D_1 = 16832$		59838.0 (5)	61827.6 (4)	52251.8 (8)	
$^5D_2 = 16412$			62246.7 (2)	52671.2 (9)	
$a = 22775$		53895 (4)	55884 (3)		
$b = 18970$			59689 (2)	50115 (9)	
$c = 16489$			62170 (0)	52593 (4)	
$d = 16054$					39235.3 (2)
$e = 15183$		61489 (2)	63476 (2)	53900 (10)	40107 (3)
$f = 13379$		63290 (2)	65280 (2)	55704 (6)	
$g = 13357$	62766 (1)	63313 (2)	—	55726 (3)	
$h = 13316$			65341 (2)	55767 (6)	

\* 'Phil. Mag.,' vol. 4, p. 486 (1927).

$mp\ ^5P—md\ ^5D$ , reference should be made to Fowler's "Report"\* and the paper by Gibbs and Ruedy (*loc. cit.*). The new levels identified in this work have been arbitrarily designated by the symbols  $a, b, c$ , etc., since any definite assignment is considered impossible with the present data. The presence of the line  $\nu 41421.2$  in the calculated position is further confirmation of the correctness of the level  $4p\ D$ . McLennan and Crawford† have referred to a level  $\nu = 55289$ , which they suggest as  $4p\ ^1S_0$ . Gibbs and Ruedy have remarked that their data neither confirm nor contradict this assignment. Table I shows that this level gives combination lines with the three levels  $5s\ ^3S_1$ ,  $4d\ ^3S_1$ , and " $c$ ." Two of these lines have been ascribed by Bloch and Bloch‡ to Se II, but our observations indicate that they belong more probably to Se I. If so, it is believed that McLennan and Crawford's identification is correct. The level  $\nu = 16054\text{ cm.}^{-1}$  is based on a single line,  $39235.3$ . The region in which the line occurs suggests that the line is due to a transition into the level  $4p\ ^1S_0$ .

Table II contains all the lines lying in the ultra-violet region which have been ascribed to Se I, together with the classifications where they have been made.

#### Se VII.

It was reported in Part I§ that the analysis of Se IV and Se V had been suggested mainly from photographs, in the region  $\lambda 1400—\lambda 400$ , of highly condensed vacuum sparks, produced by a 50 kv. transformer. The sparks were violent enough easily to excite lines up to Se VI. After the elimination of all the lines classified as due to Se IV, Se V and Se VI, there is yet an outstanding group of lines in the region  $\lambda 860—\lambda 560$ . It is difficult to judge conclusively, from experimental evidence alone, as to what stage of ionization of the atom they belong, but it is certain that they are very highly enhanced lines, occurring only under conditions when the vacuum in the spark chamber was high. It is therefore believed that nearly all of them belong to Se VII. These are entered in Table III.

Further, the spectrum of Se VII, like the arc spectrum of Ni, arises from the transition of one of the electrons of a  $3d^{10}$  configuration. By the application of the X-ray laws to the Ni-like sequence, which is known up to As VI,|| it

\* 'Phys. Soc.', 1922.

† 'Nature,' vol. 124, p. 874 (1929).

‡ 'J. Phys. Rad.,' vol. 9, p. 180 (1928).

§ K. R. Rao and Badami, 'Proc. Roy. Soc.,' A, vol. 131, p. 154 (1931).

|| Mack, Laporte and Lang, 'Phys. Rev.,' vol. 31, p. 748 (1928) and (As VI), Mack, (*unpublished results*).

is possible to predict exactly the position of the combinations  $3d^3 4s^3 D$ ,  $^1D-3d^3 4p^3 F$ ,  $^3D$ ,  $^3P$ . These fall precisely in the region where the above group of lines is present, but the observations are, however, not sufficiently

Table II.—Arc Lines of Selenium.

$\lambda$ I.A.	$\nu$ (vac.)	Classification.	$\lambda$ (vac.)	$\nu$ (vac.)	Classification.
3618.73 (2)	27626.2	$4p^1 S_0-5s^1 S_1$	1852.04 (2)	53994	
2609.47 (1)	38310.6	$^1S_0-4d^1 D_1$	†1841.11 (2)	54315	
2547.96 (2)	39235.3	$^1S_0-d$	†1838.38 (4)	54366	
2492.56 (3)	40107.3	$^1S_0-e$	†1826.30 (5)	54756	
2413.90 (8)	41421.2	$^1D_2-5s^1 S_1$	†1807.36 (8)	55330	
*2164.15 (6)	46193.2	$^3P_1-5s^1 S_2$	1795.21 (6)	55704	$4p^1 D_2-f$
2138.65 (1)	46743		1794.49 (3)	55726	$^1D_2-g$
2136.06 (1)	46790		1793.18 (6)	55767	$^1D_2-h$
*2074.75 (10)	48183.0	$^3P_2-5s^1 S_2$	1789.43 (3)	55884	$^3P_2-a$
2068.35 (1)	48332		1783.06 (3)	56083	
*2062.79 (10)	48463.2	$^3P_0-^1S_1$	*1690.72 (5)	59146.4	$^3P_0-4d^1 D_1$
2050.48 (2)	48753		1675.90 (0)	59669	
*2039.82 (15)	49008.1	$^3P_1-^1S_1$	1675.35 (2)	59689	$^3P_1-b$
			*1675.28 (5)	59691.5	$^3P_1-4d^1 D_1$
$\lambda$ (vac.)			*1671.18 (5)	59838.0	$^3P_1-^1D_2$
2026.51 (2)	49346		1666.89 (1)	59992	
2025.13 (2)	49380		1620.30 (2)	61489	$^3P_1-c$
2002.87 (2)	49928		*1621.26 (3)	61680.4	$^3P_1-4d^1 D_1$
2001.12 (1)	49973		*1617.40 (4)	61827.6	$^3P_1-4d^1 D_2$
1995.35 (9)	50115	$^1D_2-b$	1608.50 (0)	62170	$^3P_1-d$
*1990.87 (20)	50997.8	$^3P_2-5s^1 S_2$	*1606.51 (2)	62246.7	$^3P_2-4d^1 D_2$
*1919.20 (8)	52105.0	$^1D_2-4d^1 D_1$	1593.22 (1)	62766	$^3P_2-g$
*1913.81 (8)	52251.8	$^1D_2-^1D_1$	1587.44 (1)	62995	
1901.41 (4)	52593	$^1D_2-c$	1580.03 (2)	63290	$^3P_1-f$
*1898.57 (8)	52671.2	$^1D_2-4d^1 D_2$	1579.46 (2)	63313	$^3P_1-g$
1897.43 (8)	52703		1577.96 (3)	63373	
†1890.56 (1)	52922		1575.40 (2)	63476	$^3P_2-e$
†1876.49 (5)	53291		†1533.90 (2)	65193	
†1870.72 (4)	53455		1531.85 (2)	65280	$^3P_2-f$
†1869.33 (4)	53495		1530.42 (2)	65341	$^3P_2-h$
1858.90 (10)	53795		1499.83 (1)	66674	
1855.46 (4)	53895	$^3P_1-a$	1433.69 (2)	69750	
1855.29 (10)	53900	$^1D_2-e$	1413.92 (1)	70725	

\* Wave-lengths and classification are due to Gibbs and Ruedy (*loc. cit.*). Intensities are from our plates.

† These lines may belong to Se II.

‡ May be lines of sulphur.

complete to enable the writers to fit them into a suitable scheme. A re-investigation must be made before these groups can be identified. The classifications mentioned in Table III are therefore only tentative.

We are greatly indebted to Professor A. Fowler for his stimulating encouragement in our investigations.

Table III.—Highly Enhanced Lines of Selenium.

$\lambda$ vac. (int.).	$\nu$ (vac.).	Stage and classification if any.	$\lambda$ vac. (int.).	$\nu$ (vac.).	Stage and classification if any.
860.62 (3)	118195	VII $4s^2D_3-4p^2F_3$	689.96 (1)	144936	V $4d^1D_3-4f^1F_3$
864.37 (6)	117048		687.61 (2)	145431	
840.37 (0)	118995		686.48 (1)	145871	
818.95 (1)	122108		684.22 (1)	146152	
818.63 (2)	122155	VII $4s^2D_3-4p^2F_3$	680.48 (0)	148955	Probably VI.
817.48 (2)	122327		678.86 (3)	147306	
798.48 (00)	125238		674.59 (8)	148238	
779.88 (00)	128225		665.44 (1)	150277	
778.17 (4)	128507	VII $4s^2D_3-4p^2F_3$	655.12 (0)	152644	Probably VI.
774.19 (2)	129161		642.95 (2)	155533	
772.26 (3)	129490		635.80 (4)	157282	
771.59 (2)	129602		634.49 (4)	157607	
765.14 (2)	130695	VII $4s^2D_3-4p^2D_3$	626.53 (2)	159609	Probably VI.
761.98 (0)	131237		622.50 (3)	160643	
759.79 (2)	131615		613.64 (3)	162962	
756.53 (0)	132183		611.08 (3)	163645	
728.87 (1)	137199	VII $4s^2D_3-4p^2D_3$	607.19 (0)	164693	Probably VI.
720.98 (3)	138700		599.91 (2)	166692	
716.69 (1)	139530		591.28 (0)	169125	
714.08 (1)	140040		572.00 (1)	174825	
708.27 (4)	141790	VII $4s^2D_3-4p^2D_3$	563.68 (0)	177406	Probably VI.
690.35 (2)	144854		561.28 (0)	178164	

*Se I.*—About 50 lines have been newly ascribed to the neutral atom of selenium mainly by a study of uncondensed discharges through capillary tubes, and of arcs in an atmosphere of nitrogen. Eight new levels have been identified.

*Se VII.*—A study of very highly condensed and violent vacuum sparks has led to the assignment of about 40 lines in the region  $\lambda$  860– $\lambda$  560 to the spectrum of Se VII and some tentative classifications are suggested.

*The Effect of Thermal Agitation on Atomic Arrangement in Alloys.*

By W. L. BRAGG, F.R.S., and E. J. WILLIAMS, D.Sc., Manchester University.

(Received December 29, 1933.)

Forming the subject of the Bakerian Lecture by Professor W. L. Bragg, F.R.S.

(Read June 28, 1934.)

*Introduction.*

When two metals are alloyed together in various proportions, a series of solid phases is formed. A characteristic phase diagram of a binary alloy system has regions of single phase, throughout which the alloy is homogeneous, alternating with regions in which two neighbouring phases coexist. The composition of a single phase can be varied continuously over a certain range. This feature of an alloy is in contrast to the constant atomic ratio of a chemical compound, and is explained by the nature of the binding forces in an alloy which are predominantly those between the metal atoms of both kinds on the one hand and the common electronic system on the other hand, as opposed to the binding forces between atom and atom which predominate in other chemical compounds.

Not only may the atomic ratio in a given phase be varied, but also an orderly space distribution of one kind of atom relative to the other, as found in typical chemical compounds, does not necessarily exist in an alloy. Although each phase is distinguished by possessing a characteristic crystalline structure which differs from that of other phases in the same alloy system, yet this structure may be merely an orderly arrangement of sites occupied by atoms. The manner in which the atoms are distributed amongst the sites of a given phase is often variable, and is, for instance, affected by the thermal treatment which the alloy has undergone.

An alloy is a system in dynamic equilibrium. The atomic sites are determined by the interaction between metal atoms and the common electronic system, but atoms of different kinds are constantly being interchanged between one site and another owing to thermal agitation, without destroying the crystalline structure of the phase. This is shown by the rapid inter-diffusion of metals at temperatures far below the melting point and by changes in structure which can be followed by X-ray analysis. Although at room temperature the process of atomic interchange may have slowed down so much as to be

inappreciable, at some time in its history the alloy has passed through a high temperature where the process is rapid. The character of its structure has been impressed upon it at this stage, and we must in every alloy seek for a clue to the structure in the dynamic equilibrium prevailing at such temperatures. There exist, of course, all intermediate stages between such alloys and chemical compounds. In the latter both constant atomic ratio and a permanent orderly arrangement are essential features. The place of any compound between these extremes depends upon the relative importance of the metallic structure on the one hand, and the direct interatomic forces on the other hand.

The present paper is concerned with the effect of thermal conditions upon the manner of distribution of the atoms amongst the phase sites. If the metal atoms of either kind not only played essentially the same part in the structure, but were effectively identical, their distribution amongst the phase sites would be random. Actually the differences between them tend to cause atoms of the one kind to segregate into certain particular sites forming an orderly arrangement, because such an arrangement has a lower potential energy than one of disorder. Thermal agitation has the opposing effect of creating a random arrangement. The actual state of dynamical equilibrium of an alloy is one in which the two processes balance.

Tammann\* in 1919 put forward the hypothesis that in alloys subjected to long annealing the atoms segregate into regular positions, in order to explain changes in resistivity which are observed. The existence of such a segregation in gold-copper alloys was experimentally observed by Johansson and Linde† by means of X-ray analysis. When the alloy  $\text{Cu}_3\text{Au}$  is first prepared, for instance, the analysis shows that it has a face-centred cubic structure like pure gold or copper, with atoms distributed at random among the points of the lattice. After annealing for several days at  $300^\circ\text{C}$ ., additional lines appear in the X-ray photographs which indicate that the gold atoms now occupy the points of a simple cubic lattice, while the structure of phase-sites remains face-centred cubic. A face-centred cubic lattice may be regarded as composed of four identical interpenetrating simple cubic lattices, and the gold atoms segregate into one of these lattices relegating copper atoms to the remaining three, fig. 1. A more complex FeAl system has been examined in detail by Bradley and Jay.‡ The iron and aluminium atoms in all alloys of compositions between

\* 'Z. Anorg. Chem.,' vol. 107, p. 1 (1919).

† 'Ann. Physik,' vol. 78, p. 439 (1925).

‡ 'Proc. Roy. Soc.,' A, vol. 136, p. 200 (1932).



pure Fe and FeAl are arranged in a body-centred cubic phase structure, but two types of orderly arrangement of the Al atoms appear whose nature and perfection depend upon composition and thermal treatment.

Such cases, where the phase-sites considered by themselves have a very simple and symmetrical structure such that all points are geometrically equivalent, but where atoms of one kind tend to segregate into one particular set of the sites and mark out a pattern on a larger scale superimposed on the

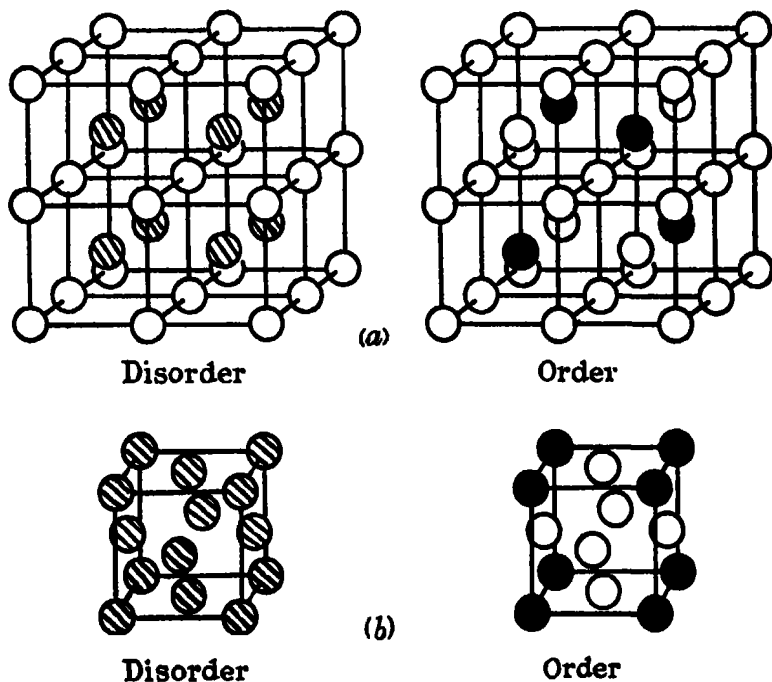


FIG. 1.—(a)  $\text{Fe}_3\text{Al}$ ,  $r = \frac{1}{2}$ ,  $n = N/2$ .     $\bullet = \text{Al}$ ,  $\circ = \text{Fe}$ .     $\otimes = 0.5 \text{ Al}, 0.5 \text{ Fe}$ .  
 (b)  $\text{Cu}_3\text{Au}$ ,  $r = \frac{1}{2}$ ,  $n = N$ .     $\bullet = \text{Au}$ ,  $\circ = \text{Cu}$ .     $\otimes = 0.25 \text{ Au}, 0.75 \text{ Cu}$ .

pattern of sites, are called Superlattices (*Überstrukturen*). Dehlinger and Graf\* in a very interesting paper have given an account of the effect of thermal treatment on the AuCu system, and have compared the X-ray results with observations of physical properties of the alloy and have discussed the theory of the effects. The general idea upon which the present paper is based, the opposing effect of temperature and of the lower potential energy of order, is that adopted by Dehlinger and others. An attempt, however, is here made to consider the kinetics of such systems more closely and also to consider the allied problem of the rate of approach to the equilibrium state when the system

\* 'Z. Physik,' vol. 64, p. 359 (1930).

is displaced from it. The latter is closely connected with the processes of annealing and quenching. It must again be emphasized that we are concerned here not with a change from one phase to another, but with continuous changes of atomic arrangement within a single phase system.

### 1. *The Relation between Degree of Order and Temperature.*

We shall first consider the way in which the degree of order of an alloy which forms a superlattice structure depends upon the temperature. It will be supposed that at each temperature the alloy is in a state of dynamical equilibrium, the rate at which atoms are being thrown out of ordered positions into disordered positions by thermal agitation being equal to the rate at which the reverse process takes place.

The degree of order of the structure may be defined as follows: We may consider the structure as formed by first placing at every point of the phase-pattern an atom of metal B, and then replacing at certain points an atom B by an atom of metal A, which will be termed an (A — B) replacement. These replacements are continually wandering through the phase-pattern. Let there be  $n$  positions in a crystal block in which such replacements may take place. The number  $n$  may not be equal to the total number  $N$  of atomic sites in the crystal block; in the  $\text{Fe}_3\text{Al}$  structure of fig. 1, for instance,  $n = N/2$ , since the replacements characteristic of this superlattice even when it is only partially formed are confined to one-half of the lattice positions (cube centres in the figure). Let there be  $m$  positions of order in the block ( $\alpha$  positions), and  $(1 - r)n$  positions of disorder ( $\beta$  positions) and further let the number of A atoms be  $rn$ . In other words, there are in all just sufficient (A-B) replacements in the crystal to make a perfect superlattice if they all occur in the correct positions. It may be seen from fig. 1 that  $r = \frac{1}{2}$  in  $\text{Fe}_3\text{Al}$ , but that in  $\text{Cu}_3\text{Au}$   $r = \frac{1}{4}$ . It must be recalled that we are dealing with cases of simple lattices where the  $\alpha$  and  $\beta$  positions are identical as regards the phase-pattern, and that we can only define the  $\alpha$  positions of order as a set distinguished by having a higher proportion of replacements than other similar sets of phase-sites.

We define the degree of order  $S$  as follows: Let  $p$  be the probability that an  $\alpha$  position is occupied by a replacement. Then

$$S = \frac{\text{Actual value of } p - \text{value of } p \text{ for complete disorder}}{\text{Value of } p \text{ for complete order} - \text{value of } p \text{ for complete disorder}},$$

$$= \frac{p - r}{1 - r}.$$

When order is complete,  $p$  is unity and  $S$  is therefore unity. When disorder is complete and the arrangement of the replacements quite random, only a fraction  $r$  of the replacements will be in the positions of order. Therefore  $p = r$ , and  $S = 0$ .

Consider the crystal at a given temperature  $T$ , and with a given degree of order  $S$  (which need not be the equilibrium degree of order at  $T$ ). We define a quantity  $V$ , for this  $S$  and  $T$ , as the increase in the potential energy of the crystal when one replacement is moved from an  $\alpha$  (ordered) to a  $\beta$  (disordered) position, all the other replacements being unaltered. We may write

$$V = V(S, T).$$

It is tacitly assumed that all  $\alpha$  positions are identical and all  $\beta$  positions identical, or that we may regard  $V$  as an effective average value although there are fluctuations in the local degree of order at different points.

Suppose again the crystal to be at a given temperature  $T$  and a replacement to involve an energy  $V$ . The condition that a degree of order  $S$  is one of dynamical equilibrium for this particular  $V$  leads to a relation

$$S = S(V, T).$$

If now these two relations for a given temperature  $T$  are plotted with  $V$  and  $S$  as co-ordinates, the intersection of the curves gives equilibrium values of  $S$

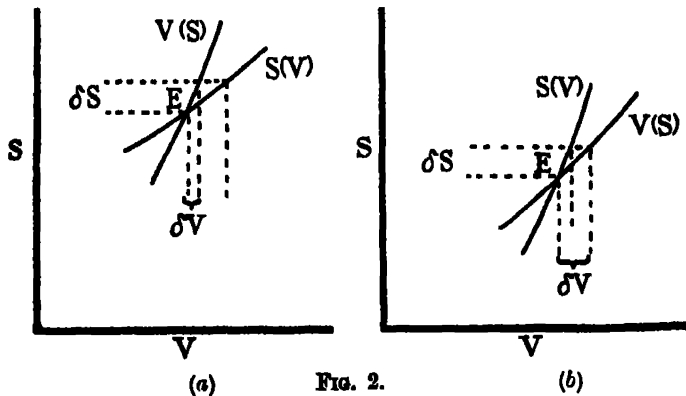


FIG. 2.

and  $V$  for that temperature. In other words, the equilibrium degree of order  $S_e$  creates a value of  $V_e$ , the potential energy  $V$  involved in an atomic interchange, which is such that it makes the number of transitions from order to disorder equal to the number from disorder to order, and thus preserves  $S_e$ .

Fig. 2 shows that such intersections may correspond to stable or unstable equilibrium. Fig. 2a represents stable equilibrium,  $V(S)$  having a greater

slope than  $S(V)$ . If  $S$  is for any reason increased a little above  $S_e$ , for instance,  $V$  undergoes a slight increment  $\delta V$  given by the  $V(S)$  curve. It will be clear from the  $S(V)$  curve that this increment is insufficient to maintain the higher degree of order, which therefore decreases towards the equilibrium value. On the other hand, in fig. 2b the increase in  $V$  produced by an increase in  $S$  is higher than that required to maintain the new degree of order.  $S$  therefore increases still further and the structure is in unstable equilibrium.

We may now consider the general nature of the  $V(S, T)$  and  $S(V, T)$  relations. A discussion based on more detailed assumptions follows in the next paragraph.

At a given temperature, the degree of order increases with increase of  $V$ , being zero for  $V = 0$  and unity when  $V$  becomes very great. Also, for a given  $V$ , it decreases with increasing temperature being unity for  $T = 0$  and approach-

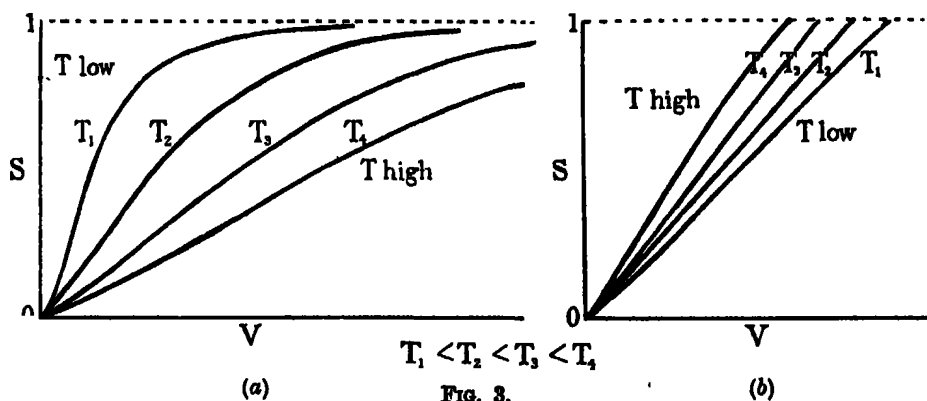


FIG. 3.

ing zero as  $T$  becomes very great. A family of  $S(V, T)$  curves for different temperatures will therefore have the character shown in fig. 3a. It may indeed be anticipated that  $S(V, T)$  has the form  $S(V/kT)$  as in similar thermodynamical relationships, so that the curves for increasing temperatures are similar but for an increase in the scale of the abscissa  $V$ .

$V(S, T)$  must vanish at  $S = 0$ , because the criterion between positions of order and disorder ceases to exist. It is on this assumption that our discussion is based. It is justified by the peculiar nature of the alloy superlattice, where the positions of order are not distinguished in any way from those of disorder except by their having been singled out for a partial segregation of atoms of one kind.  $V(S, T)$  must rise to a maximum value  $V_0$  at  $S = 1$ , since in this case a replacement moving from an ordered to a disordered position is subjected to the full influence of a complete set of ordered neighbours. It is to be expected that  $V$  depends mainly upon  $S$ , being almost independent of

the temperature.\* Further, an approximation to a linear relation between  $V$  and  $S$  is probable, since the excess number of ordered over disordered neighbours influencing a given (A — B) replacement is on the average proportional to  $S$ . Fig. 3b shows the general nature of the  $V(S, T)$  curves.

The equilibrium degree of order at a temperature  $T$  is given by the intersection of the two corresponding curves in fig. 3. If we picture the successive intersections, it will be seen that the variation of the degree of order with temperature has the following characteristics. At high temperatures, the only intersection is at 0, corresponding to complete disorder, and this state is stable because the  $V(S)$  curve has a greater slope than the  $S(V)$  curve. At low temperatures the intersection at 0 represents unstable equilibrium, and a second intersection at a finite value of  $S$  represents stable equilibrium. This value of  $S$  tends to unity as  $T$  approaches zero. This second intersection only exists below a certain critical temperature  $T_c$ , above which the structure is always in complete disorder. Further, from the nature of the curves, it will be clear that the point of intersection rises very rapidly from zero towards unity as  $T$  falls below the critical temperature. There is therefore on cooling a sudden onset of order below  $T_c$ , followed by a more gradual increase towards unity.

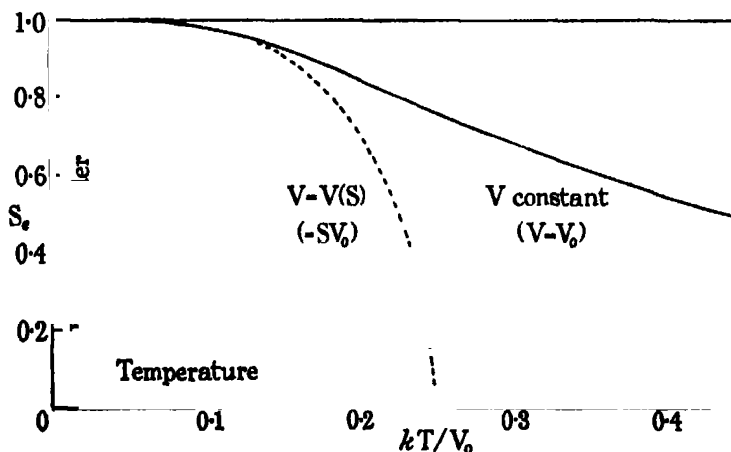


FIG. 4.

This characteristic relation is the result of our general assumption that  $V$  depends upon  $S$ , and tends to zero as  $S$  tends to zero. If  $V$  were constant, there would be a gradual falling off of the degree of order with increasing temperature, as in fig. 4. According to our theory, the decrease in  $S$  involves

\*  $V$  may decrease slightly with increasing  $T$ , in correspondence with a decrease of elasticity.

a decrease in  $V$ , and hence the degree of order falls off more steeply and reaches zero at the critical temperature as shown by the dotted curve.

The formal relation leads to solutions of two types, illustrated in fig. 5. If the  $V(S)$  curve is more convex to the  $V$  axis than the  $S(V)$  curve (fig. 5a), there can only be one intersection  $E$  in addition to the intersection at 0. The relation between  $S_e$  and  $T$  is in this case of the form shown in fig. 5b. On the other hand, if  $V(S)$  is less convex to the  $V$  axis than  $S(V)$ , there will exist in a certain temperature range a triple intersection as shown in fig. 5c. Both  $O$  and  $E_2$  represent stable equilibrium,  $E_1$  unstable equilibrium. The

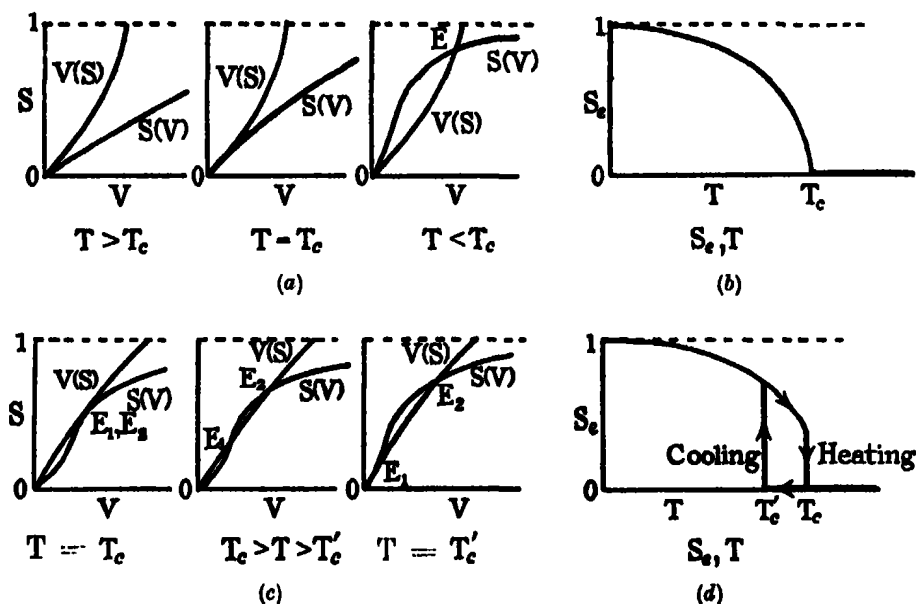


FIG. 5.

corresponding relation between  $S_e$  and  $T$  has a hysteresis loop as shown in fig. 5d. If we start with the alloy at a low temperature and gradually raise the temperature, the intersection  $E_2$  will move along the curve until it coalesces with  $E_1$ . Past this point it ceases to exist and the degree of order falls abruptly to zero. We will call this temperature  $T_c$ . If the alloy is now cooled, starting in the state of complete disorder represented by the intersection at  $O$ , this remains a state of stable equilibrium (by our criterion) as long as the  $V(S)$  curve has a greater slope at the origin than the  $S(V)$  curve. It is still stable, therefore, even when there is a further stable intersection at  $E_2$  and remains so till a temperature  $T'_c$ , which is lower than  $T_c$ . At this point order rises from zero to a finite value represented by the appropriate upper intersection.

There will, therefore, be two critical temperatures at which the degree of order alters discontinuously from zero to a finite value or *vice versa*. When the alloy is being warmed the critical temperature is higher than when it is being cooled.

The interpretation of this formal solution, however, requires careful examination. Of the two equilibrium states given by our theory in the above case, the one with lower thermodynamical potential is the permanently stable one. Let this be A, and suppose the crystal to be in the other state B, which is metastable. Then local fluctuations in the atomic arrangement in B which simulate A will be likely to grow and cause the whole crystal to change over into state A. On the other hand, if the crystal is in state A fluctuations simulating B will have a tendency to decrease. The hysteresis loop will therefore not be realized in practice unless the rate of cooling or heating between  $T_c$  and  $T'_c$  is so rapid that the probability of fluctuations in B which precipitate A is small in the time interval between  $T_c$  and  $T'_c$ .

In a later paragraph, we discuss the "time of relaxation  $\tau$ " of the alloy at different temperatures. When the alloy is in a state not that of equilibrium, the departure from equilibrium is reduced in a time  $\tau$  to  $1/e$  of its initial value. A simple calculation shows that during the time  $\tau$ , fluctuations in the atomic arrangement of B which simulate A will have appeared at intervals apart of a small number of lattice-spacings in the crystal. The crystal will therefore pass over into state A in this time in spite of the formal conclusion drawn from the curves that B is a stable state.

A difference between the critical temperatures for cooling and heating (irreversible change) is actually observed. We must attribute this, however, not directly to the existence of a triple intersection as in fig. 5, but to the general feature that an alloy, the temperature of which is changing, always lags behind its equilibrium state owing to the existence of the time of relaxation. It will be shown further that relaxation is particularly slow in the neighbourhood of the critical temperature, and that it may not be feasible to heat or cool so slowly as to avoid a considerable hysteresis loop, even though there is ideally only one equilibrium state at each temperature.

The general conclusion that order sets in abruptly below a critical temperature  $T_c$  has a close analogy in ferromagnetism, and there are, in fact, many points of similarity between the present treatment and the classical equations of Langevin and Weiss. We may compare the degree of order in the alloy with the intrinsic field of a ferromagnetic. The average orientation of the atomic magnets corresponds to the degree of order in the alloy, and the differ-

ence in potential energy for the parallel and antiparallel positions to the  $V$  we have considered above. There is a similar dependence of the one on the other, and on the temperature, in both cases. Hence the alloy has a critical temperature above which no order or superlattice exists, just as the ferromagnetic has a "Curie point."

The sudden onset of order below  $T_c$  will cause the curves representing the variation with temperature of physical properties of the alloy, such as resistivity, specific heat, and thermal expansion, to have a sharp inflexion at this point. These inflexions will simulate the effects of a phase change, though we are not concerned here with a phase change but with a continuous change within the one phase structure.

It will be noted that the definition of order in this treatment applies to an organized order throughout crystal blocks. We are indebted to Dr. Peierls and Dr. Bethe for pointing out that order when defined in an alternative way does not cease abruptly at a critical temperature. If we define order by considering the nature of the neighbouring atoms around a given atom of either kind in the structure, complete order means that these neighbours are in the relation of a perfect superlattice, and complete disorder that they are atoms of either kind in a random way. The two definitions lead to the same result when order is nearly complete. On the other hand, complete disorder as defined in this alternative way is never reached at finite temperatures. Even when the temperature is very high, there remains a certain probability that the immediate neighbours of an A atom are B atoms rather than A atoms and *vice versa*, if we take as an example the system where A and B atoms are equal in number. The more we take into account distant neighbours in our definition of order, the more steeply does the curve of order plotted against temperature fall towards zero and the more definite is the point which we may take to be the "critical temperature." The formation of sharp superlattice lines in the X-ray photograph, and the influence of order upon resistivity, depend upon order over large distances and may be expected to show the effects predicted on the present theory. For the specific heat, however, which depends for the greater part upon the forces between nearest neighbours, the alternative definition is more appropriate. It can be shown that the two definitions lead to variations in the specific heat of the same order of magnitude, but with a different dependence upon temperature.



## 2. Further Discussion of the $S(V, T)$ and $V(S, T)$ Curves, and their Solutions $S_s(T)$ .

To repeat the statement of the problem, there are  $rn$  (A-B) replacements distributed between  $n$  sites, of which  $rn$  are positions of order ( $\alpha$ ) and  $(1-r)n$  positions of disorder ( $\beta$ ).  $V$  is the increase in potential energy due to one replacement moving from an ordered to a disordered position. The degree of order  $S$  is defined as  $(p-r)/(1-r)$  where  $p$  is the probability that any particular replacement is in an ordered position.

Any such replacement spends part of its time in an  $\alpha$  position and part in a  $\beta$  position. At a given moment  $mp$   $\alpha$  positions will be occupied by other replacements, and so also will  $rn(1-p)$   $\beta$  positions. The replacement we are considering will thus have open to it  $(rn-mp)$   $\alpha$  positions, and  $\{(1-r)n-rn(1-p)\}$   $\beta$  positions. By Boltzmann's relation

$$\begin{aligned}\frac{p}{1-p} &= \frac{\text{Number of available } \alpha \text{ positions}}{\text{Number of available } \beta \text{ positions}} \cdot e^{V/kT} \\ &= \frac{r(1-p)n}{(1-2r+rp)n} \cdot e^x,\end{aligned}\quad (5)^*$$

where  $x = V/kT$ .

Solving this equation for  $p$ , and substituting in  $S = (p-r)/(1-r)$  we find

$$S(V, T) = 1 - \frac{\{4r(1-r)(e^x - 1) + 1\}^{\frac{1}{2}} - 1}{2r(1-r)(e^x - 1)}.\quad (6)$$

The curves for  $S$  as a function of  $x$  are shown in fig. 6, for the cases  $r = \frac{1}{2}$  and  $r = \frac{1}{4}$ . It will be seen that the curves are very similar, except in the neighbourhood of the origin. For small  $x$ , retaining terms in  $x^2$ , the relation becomes

$$S(x) = r(1-r)x + \frac{1}{2}r(1-r)(1-2r)^2x^2.\quad (7)$$

The curve therefore starts from the origin with a finite slope  $r(1-r)$  which is indicated by dotted lines, and the coefficient of  $x^2$  is zero for  $r = \frac{1}{2}$ , and positive for any other case (i.e., the curve is in general convex to the  $x$  axis near the origin).

In the case  $r = \frac{1}{2}$  the relation (6) simplifies to

$$\begin{aligned}S(V, T) &= 1 - 2(e^{x/2} - 1)/(e^x - 1) \\ &= (1 - e^{-x/2})/(1 + e^{-x/2}) \\ &= \tanh(x/4), \quad x = V/kT.\end{aligned}\quad (8)$$

\* A derivation of this formula based on the law of Mass Action is indicated in § 6.

This is the case where there are equal numbers of ordered and disordered positions, as in the  $\text{Fe}_3\text{Al}$  structure shown in fig. 1.

Passing now to the dependence of  $V$  upon  $S$  and  $T$ , we have seen that  $V(S, T)$  is zero for  $S = 0$  and a maximum  $V_0$  for  $S = 1$ , and that it is not likely to be appreciably dependent upon temperature. It will be assumed that the relation is linear

$$V = V_0 S, \quad dV_0/dT = 0. \quad (9)$$

The equilibrium value of  $S$  at a given temperature  $T$  is one which satisfies both (6) and (9).

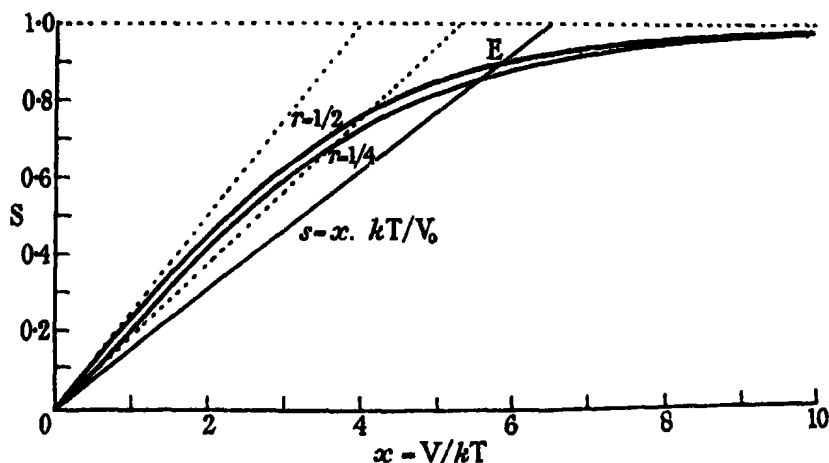


FIG. 6.

To illustrate the nature of the solution, it will be derived for the case  $r = \frac{1}{2}$  where it has been seen that equation (6) reduces to the simpler form given in (8). We have

$$S_e = \tanh(V/4kT) = \tanh(x/4) \\ V = V_0 S_e.$$

To solve this, we plot, as in fig. 6, the curve  $S = \tanh(x/4)$  and a straight line  $S = xkT/V_0$  for each temperature. The point where the straight line cuts the curve gives the degree of order  $S_e$  characteristic of the temperature  $T$ .

The relation  $S = \tanh(x/4)$  approximates to  $S = x/4$  when  $S$  and  $x$  are small. Hence the line  $V = V_0 S$  is tangential to the curve at the origin at a critical temperature  $T_c$  such that

$$4kT_c/V_0 = 1 \\ T_c = V_0/4k. \quad (10)$$

Above  $T_c$  the only state of equilibrium is one of complete disorder. Below it complete disorder becomes unstable, and the other point of intersection rapidly rises towards  $S = 1$  as the temperature is lowered. We have

$$\begin{aligned} S_c &= \tanh(V/4kT) \\ &= \tanh(V_0 S_c/4kT) \\ &= \tanh(S_c T_c/T). \end{aligned} \quad (11)$$

Hence  $S_c$  is a function of  $T/T_c$  alone. It is plotted against  $T/T_c$  in fig. 7.

$S_c$  is zero for  $T > T_c$ . At  $T = T_c$  there is a sharp inflexion and the curve runs vertically at first; it then bends over so as to approach  $S = 1$  asymptotically. When  $T$  is slightly less than  $T_c$ , the relation  $S_c = \tanh(S_c T_c/T)$  can be shown to reduce to

$$S_c^2 = 3(T_c - T)/T_c, \quad (12)$$

which is a close approximation as far as  $S_c = 1/2$ .

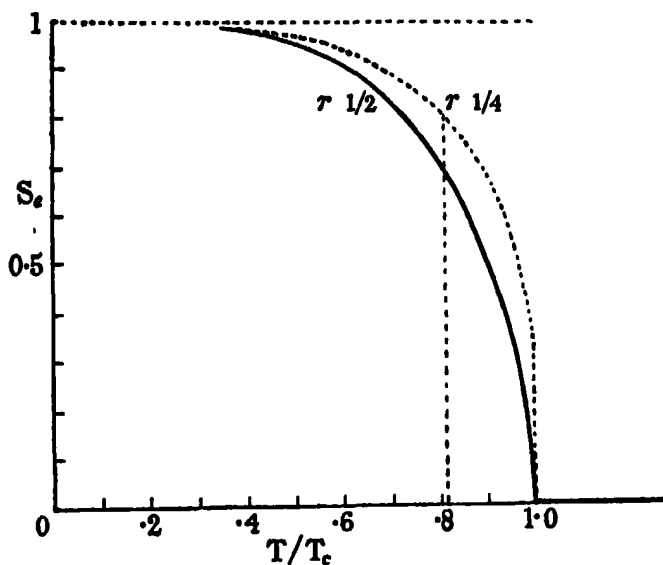


FIG. 7.— $r = \frac{1}{2}$ .

$T/T_c$	$S_c$	$T/T_c$	$S_c$
1	0	0.85	0.61
0.98	0.25	0.80	0.71
0.96	0.33	0.70	0.83
0.94	0.39	0.60	0.90
0.92	0.45	0.50	0.95
0.90	0.50	0.40	0.98

The relation (10) gives the value of  $V_0$  for a given value of  $T_c$ . For instance, the critical temperature at which the formation of the  $\text{Fe}_3\text{Al}$  superlattice sets in is  $560^\circ\text{C}$ . ( $833^\circ\text{K}$ .) according to Sykes. The corresponding value of  $V_0$  is 0.29 electron-volts. This must only be considered as an estimate of the order of  $V_0$ , in view of the sensitivity of the calculated value of  $T_c$  to small alterations in the form of the  $V(S, T)$  and  $S(V, T)$  relationships.

The form of the relation between  $S_c$  and  $T$  for  $r = \frac{1}{2}$  can be calculated in a similar way. As will be clear from fig. 6, the corresponding curve shows an hysteresis loop because the  $S(V, T)$  curve is slightly concave upwards at the origin. Its form is shown by the dotted lines; in this case  $kT_c/V_0 = 0.21$ . We do not desire, however, to attach an unwarranted weight to details of the above calculations. The commencement of order, in particular, depends in a very sensitive way on the form of the  $V(S)$  and  $S(V)$  curves near the origin. We may only claim that the theory gives a general idea of the way in which  $S_c$  varies with  $T$ , the outstanding feature of the variation being the rapid initial increase in the degree of order at a critical temperature.

### 3. *The Specific Heat of the Alloy.*

Above the critical temperature where disorder is complete, the specific heat of the alloy will be simply that due to the variation of energy of thermal agitation with temperature. Below the critical temperature, any temperature change involves a change in the degree of order  $S$ , and so in the extent to which atoms have segregated into positions of lower potential energy. An increase in order implies an evolution of heat, hence as the alloy cools there will be a sharp increase in its specific heat at the critical temperature. When times taken for the alloy to cool through equal temperature intervals are plotted against temperature, for instance, these times will suddenly increase on passing  $T_c$ .

Let us suppose the degree of order to change slightly from  $S$  to  $S + dS$  owing to a decrease in temperature  $-dT$ . The number of atoms in positions of order changes from  $np$  to  $n(p + dp)$ . Since  $S = (p - r)/(1 - r)$ ,  $dp = (1 - r) dS$ . Each atomic interchange decreases the potential energy of the whole structure by  $V = V_0 S$ . Hence the evolution of energy due to increase in order is given by

$$dE = V_0 nr (1 - r) S dS. \quad (13)$$

The evolution of energy between two states of order  $S_1$  and  $S_2$  is therefore

$V_0 nr (1 - r) (S_2^2 - S_1^2)/2$ , and the total amount for the change from disorder to order is

$$V_0 nr (1 - r)/2. \quad (14)$$

For  $r = \frac{1}{2}$ ,  $V_0 = 4kT_c$ , and (14) reduces to  $(n/2N) \cdot RT_c$  per gram equivalent.

Tammann and Heusler\* have examined what they term the "Heat of Dissociation," which according to our theory is the evolution of energy due to atoms passing into ordered positions, when  $\beta$  brass, CuZn, cools from a series of temperatures down to room temperature. Comparative measurements were made by allowing similar blocks of copper and brass to cool from the required temperature, and determining by thermo-elements the difference in temperature between the blocks during the cooling. The curve summarizing their results (p. 352, *loc. cit.*) is shown in fig. 8a. The heat evolution is a maximum

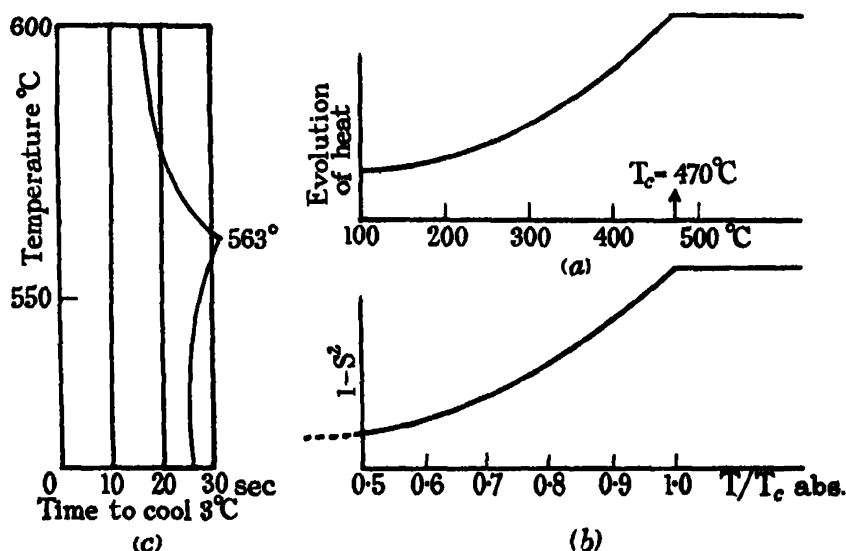


FIG. 8.—(a) and (b), CuZn,  $r = \frac{1}{2}$ . (a) Experimental, Tammann; (b) theory.  
(c) Fe<sub>3</sub>Al, (27.5% Al). Experimental, Sykes.

when the alloy cools from above 470° C., which is the critical temperature for  $\beta$  brass, and falls away as the initial temperature is lowered. According to the above calculation, the heat evolved when passing from a state of order  $S_2^2$  to a low temperature (for which  $S_2 = 1$ ) is proportional to  $(1 - S_2^2)$ . In fig. 8b  $(1 - S_2^2)$  has been plotted against the critical temperature, the ordinates being on a scale which corresponds to Tammann and Heusler's arbitrary

\* 'Z. anorg. Chem.,' vol. 128, p. 349 (1923).

vertical scale of "Heat of Dissociation." The temperature scale  $T/T_c$  is adjusted so that the absolute zero points and critical temperatures of the two curves correspond. The agreement between the experimental and theoretical curves is very close.

To obtain an estimate of the influence of this change in potential energy upon the specific heat, we may further consider the simple case for which  $r = \frac{1}{2}$ . It has been seen (equation (12)) that just below the critical temperature the degree of order is given by

$$S_c^2 = 3 (T_c - T)/T_c,$$

whence

$$= -3dT/T_c.$$

Thus, the evolution of energy for a decrease in temperature  $-dT$  in this region is by (13)

$$dE = V_0 \cdot \frac{n}{2} \cdot \frac{1}{2} \cdot \frac{3}{2} \frac{dT}{T_c}.$$

Or, since  $T_c$  is here equal to  $V_0/4k$

$$dE = 3nkdT/2.$$

During the same temperature interval, the decrease  $dE_0$  in energy of thermal movement is  $3NkdT$ , assuming Dulong and Petit's law to hold,  $N$  being the total number of atoms in the crystal block. Thus  $dE/dE_0 = n/2N$ . In other words, the specific heat should increase abruptly in passing from above to below the critical temperature, in the simple ratio

$$1 + n/2N : 1. \quad (15)$$

For the  $\text{Fe}_3\text{Al}$  structure,  $n/N$  is  $\frac{1}{2}$ , and therefore the specific heat should increase by 25%.

As in previous calculations, this can only be regarded as an estimate of the order of the effect, since the precise way in which order commences cannot be clearly predicted. If the relation between the  $V(S, T)$  and  $S(V, T)$  curve is that discussed in case II above, order passes discontinuously from zero to a finite value at the critical temperature. In this case there will be an effective latent heat associated with the change, and a curve of inverse cooling rate plotted against temperature will show a "thermal arrest point." It is interesting to note, however, that the order of the effect agrees with what is observed experimentally. For instance, the time taken for the  $\text{Fe}_3\text{Al}$  alloy to cool equal temperature intervals are plotted horizontally against temperature vertically in fig. 8c, the results being due to Sykes. At the critical temperature of  $560^\circ$  the

curve shows a small cusp representing an increase in specific heat of 30% to 40%.

#### 4. The Resistivity of the Alloy.

It will be assumed that the alloy with a perfect superlattice structure has a resistance like that of a pure metal. When  $S = 1$ , its resistivity is given by

$$\rho_1 = \alpha_1 T.$$

On the other hand, the resistivity of an alloy in a state of complete disorder does not fall to zero as the temperature approaches zero, because resistance is caused by the disorder just as it is by thermal agitation. We assume that when  $S = 0$

$$\rho_0 = \rho_d + \alpha_0 T.$$

If we may assume the resistivity of the alloy in a state of order  $S$  to have an intermediate value it will be given by

$$\begin{aligned} \rho_s &= S\rho_1 + (1 - S)\rho_0 \\ &= \rho_d(1 - S) + T\{\alpha_1 S + \alpha_0(1 - S)\}. \end{aligned} \quad (16)$$

Since  $S_c$  is zero above the critical temperature, and increases rapidly below it, a curve in which resistivity is plotted against temperature should show a very rapid downward inflexion just below the critical temperature.

We may compare the theoretical relation given by (16) with experimental measurements of the resistivity of gold-copper alloys given in a highly interesting paper by Kurnakow and Ageew.\* Curves for alloys of composition AuCu and AuCu<sub>3</sub> are shown in figs. 9 and 10. The full-line curves are plotted from tables given in their paper, the resistance of the alloy being measured as it cooled from above the critical temperature to room temperature. In fig. 9 the curves for Au, Cu, and AuCu are shown. The precise rate of cooling is not defined, but it is stated that the process lasted 12-14 hours. The dotted curve is that given by equation (16), the critical temperature  $T_c$  here being 460° C. (733° K.). The dependence of  $S_c$  on  $T$  is taken to be that given by  $r = \frac{1}{2}$  in fig. 7. Experimental and theoretical curves are similar, but it appears that the AuCu structure does not reach the state of complete order. This effect is still more marked in fig. 10. The lower experimental curve is that for a total time of cooling of 12-14 hours, while the upper curve is that for an alloy cooled in 3 hours. The large difference in the curves for the two rates of cooling supports

\* 'J. Inst. Metals,' vol. 46, p. 481 (1931).

the view that the final degree of order is far from complete in both cases, and indicates that the curve for resistivity would be not dissimilar to the theoretical curve if—as is assumed in calculating the latter—the alloy had been in equilibrium at each temperature.

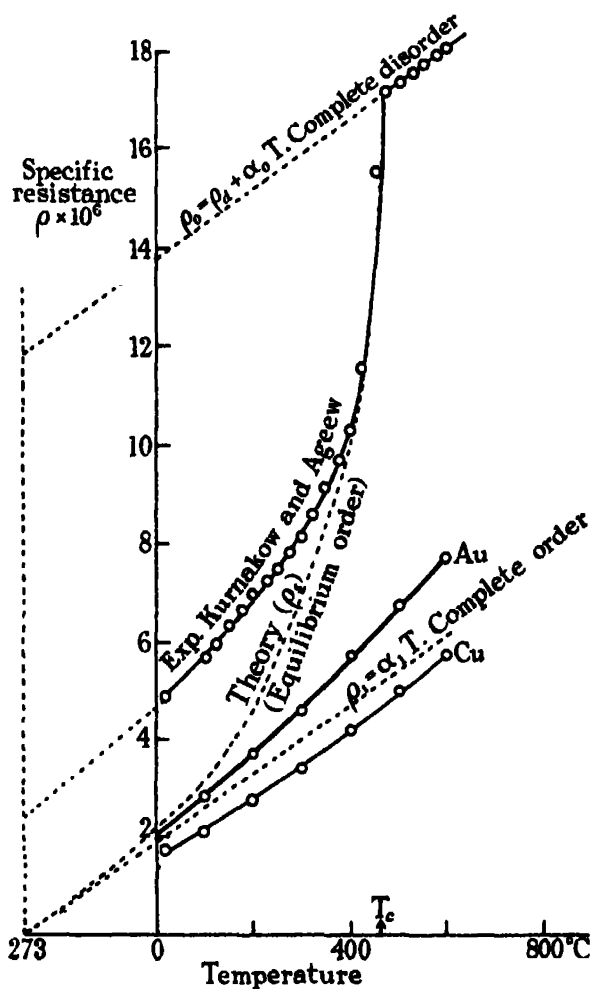


FIG. 9.—AuCu.

The influence of the rate of cooling upon the form of the curves properly belongs to the next section, where further reference will be made to it.

The simple formula  $S_e = \tanh (S_e T_e / T)$  for the dependence of equilibrium order upon temperature used above does not strictly apply to either of the cases considered. In  $\text{AuCu}_2$ ,  $r = \frac{1}{2}$  and the curve for  $S_e$  departs somewhat from the form given by  $S_e = \tanh (S_e T_e / T)$ . In  $\text{AuCu}$  the use of this formula is still more



open to objection because the change from disorder to order is accompanied by a change from cubic to tetragonal symmetry which is outside the scope of the present theory. We therefore do not wish to emphasize points of detail, and only advocate the general agreement between the observed resistivity curve and that to be expected from the type of relation between degree of order and temperature which our theory predicts.

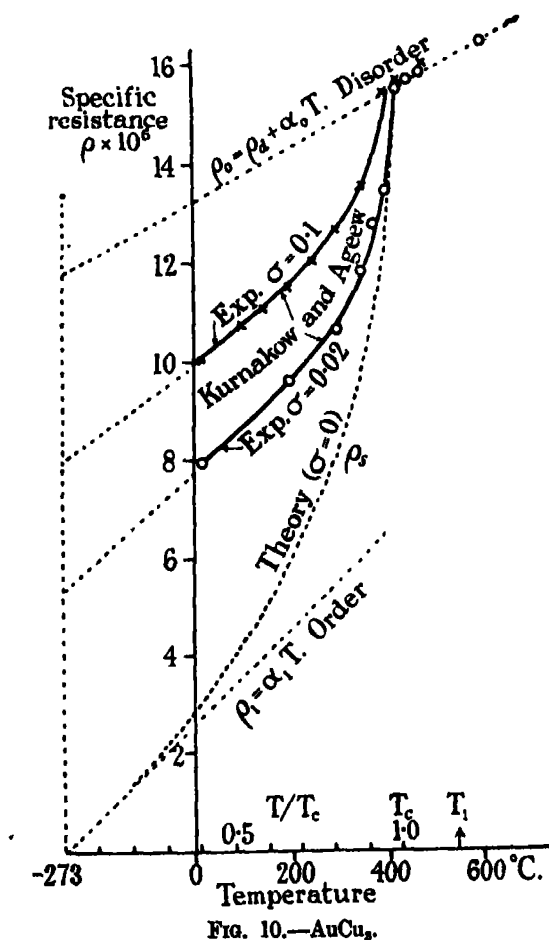


FIG. 10.—AuCu<sub>3</sub>.

### 5. The Rate of Approach to the Equilibrium State.

It has so far been assumed that the alloy takes up at each temperature the equilibrium state characteristic of that temperature. The difference between annealed and quenched alloys, and the effect of different rates of cooling on the resistivity, show that this is not so. It must be supposed that a potential

barrier has to be surmounted before each interchange of atomic position can take place, and that this only happens when the thermal movements conspire to concentrate at the required point a sufficient kinetic energy of the atoms concerned, with a suitable orientation of their momenta. At higher temperatures such interchanges are frequent, but at lower temperatures the localization of sufficient energy becomes an increasingly rare event.

We will consider an alloy whose state of order  $S$  is slightly different to that of dynamical equilibrium. This difference can be defined by saying that the alloy is at a temperature  $T$ , but that its state is characteristic of a slightly different temperature  $\theta$ . The alloy approaches the state of equilibrium exponentially according to the equation

$$d\theta/dt = -(\theta - T)/\tau. \quad (17)$$

In this equation  $\tau$  is the "time of relaxation" of the alloy or the time taken for the departure from equilibrium to be reduced to  $1/e$  of its initial value.

When the alloy is in stable dynamical equilibrium, as many atoms are passing per second from ordered to disordered positions as in the reverse direction. When not in equilibrium this balance is disturbed in such a direction that the alloy relaxes towards the equilibrium state. The rate of return is calculated in the next section. The results can be anticipated here by saying that  $\tau$  depends upon the temperature of the alloy, as well as upon certain factors practically independent of temperature, according to the equation

$$\tau = Ae^{W/kT}. \quad (18)$$

Here  $A$  is effectively a constant, and  $W$  is the "activation energy" required for the interchange of atomic position. It will further be shown in the next section that an estimate of the order of the constant  $A$  can be made.  $A$  proves to be about  $10^{-12 \pm 2}$ ; only its order of magnitude is required in the present calculation.

The characteristic physical quantities in this relation are  $A$  and  $W$ . Rather than use  $A$  and  $W$  it is more convenient to describe the dependence of the time of relaxation upon temperature in terms of  $A$  and a characteristic temperature  $T_1$ , which we define as the temperature at which the time of relaxation of the alloy is 1 second, i.e.,

$$Ae^{W/kT_1} = 1.$$

Substituting for  $W$  in (18), we obtain

$$\tau = e^{-\log A \left( \frac{T_1}{T} - 1 \right)}.$$

The way in which  $\tau$  depends upon temperature is shown in Table II. In this table,  $\log_{10} A$  has been given the value  $-12$ . Alternative values in the neighbourhood of 12 necessitate an alteration in the  $T/T_1$  scale, but it is clear from the nature of the relation that this alteration is not large.

Table II.

$T/T_1$ (K).	$\tau$ (in seconds).
1.2	0.01
1.1	0.08
1.0	1
0.9	20
0.8	$10^2$ (17 minutes)
0.7	$1.3 \times 10^4$ (36 hours)
0.6	$10^8$ (3 years)
0.5	$10^{13}$ (30,000 years)

Such a table enables us to see what will be the result of a specified thermal treatment of an alloy. Let us suppose, for instance, that  $T_1$  is  $500^\circ \text{C}$ . ( $773^\circ \text{K}$ ). In order to bring the alloy into a state of equilibrium characteristic of  $200^\circ \text{C}$ . ( $T/T_1 = 0.6$ ) it must be annealed for many years. Several days suffice to bring it to equilibrium at  $270^\circ \text{C}$ ., and an hour to bring it to equilibrium at  $350^\circ \text{C}$ . On the other hand, states characteristic of temperatures above  $650^\circ \text{C}$ . ( $T/T_1 = 1.2$ ) relax so rapidly that they cannot be preserved by quenching.

The result of thermal treatment can be defined more closely in the following way. Let it be supposed that the alloy is being cooled at a uniform rate of  $\sigma$  degrees per second from a high temperature. Since equilibrium is not immediately attained, the state of the alloy when the temperature has fallen to  $T$  is one characteristic of a higher temperature  $\theta$ . At high temperatures the lag ( $\theta - T$ ) is small because the time of relaxation is so short, but as  $T$  diminishes ( $\theta - T$ ) increases. Finally, the alloy relaxes so slowly that it is effectively "frozen" in a state corresponding to a temperature  $\theta_0$  even when  $T$  has come down to zero. We will examine the dependence of the final temperature  $\theta_0$  upon the rate of cooling  $\sigma$ .

At a given instant, the alloy is changing its state according to the equation

$$d\theta/dt = -(\theta - T)/\tau.$$

The temperature  $T$  is simultaneously decreasing at the rate

$$dT/dt = -\sigma,$$

hence the relation between  $\theta$  and  $T$  is given by the differential equation

$$d\theta/dT = (\theta - T)/\sigma\tau \quad (19A)$$

i.e.,

$$\theta - T = \sigma\tau (d\theta/dT). \quad (19B)$$

The type of curve given by this differential equation is shown in fig. 11, where  $\theta$  is ordinate and  $T$  abscissa.  $\tau$  decreases rapidly with increasing tempera-

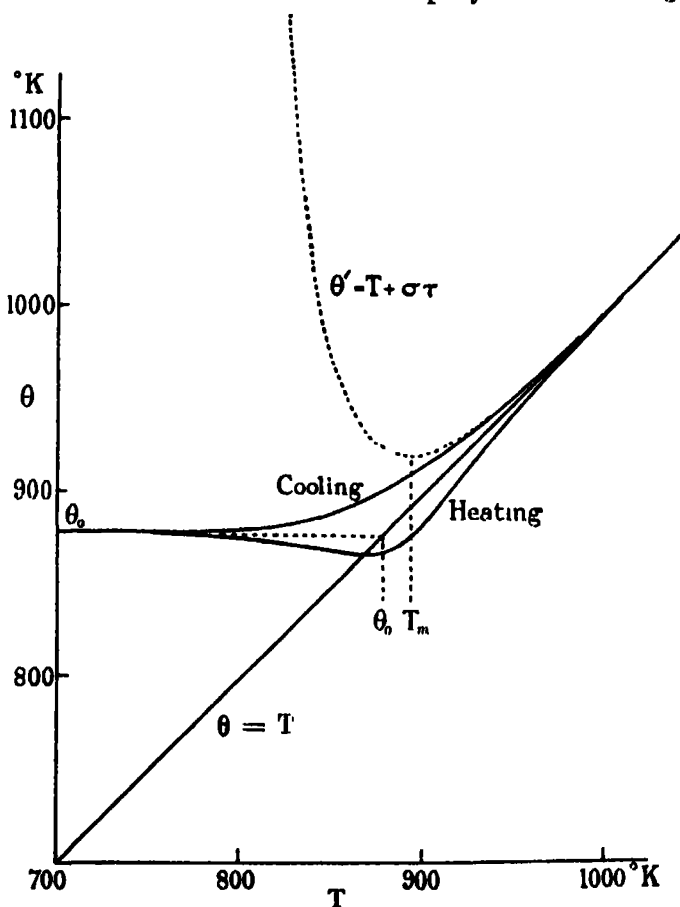


FIG. 11.

ture so that at high temperatures the right-hand side of (19B) is very small. At high temperatures the curve therefore has an asymptote  $\theta = T$ . At lower temperatures, even though  $(\theta - T)$  is appreciable, the slope of the curve is practically the same as that of the straight line  $\theta = T$ . Thus  $d\theta/dT \sim 1$ , and the curve is closely given by the relation

$$\theta - T = \sigma\tau. \quad (20)$$

This approximation in turn ceases to hold. The curve flattens out, and becomes practically horizontal at a temperature  $\theta_0$  which characterizes the final metastable state of the alloy. This section of the curve may be constructed graphically for any given alloy. It may be shown analytically, if we assume the formula (18) for  $\tau$ , that the total decrease in  $\theta$  after (20) ceases to hold is very small.

The actual curve in the figure is plotted for a rate of cooling of  $1^\circ$  per second, and a characteristic temperature  $T_1$  (for which  $\tau = 1$ ) of  $1000^\circ$  K. The final temperature  $\theta_0$  is  $880^\circ$  K. The variation of time of relaxation with temperature given by (18) was assumed.

It is interesting to note that if the alloy is heated again at the same rate as it was cooled, it will not retrace its path. It is evident that  $\theta$  will continue to decrease until it gets below the actual temperature  $T$ . This will happen in the neighbourhood of the temperature where the cooling curve flattens out. Above this temperature  $\theta$  will be given approximately by (20) but with  $\sigma$  now negative. The difference between  $\theta$  for heating and cooling gives rise to the hysteresis effect, to which previous reference has been made. Curves for both cooling and heating are shown in fig. 11.

The dotted curve in the figure represents  $\theta' = T + \sigma\tau$  which we have seen to be a good approximation to  $\theta$  as long as  $d\theta'/dT \sim 1$ . Owing to the rapid increase in  $\tau$  as the temperature decreases,  $\theta'$  has a minimum and then ascends rapidly. The ascending part of the curve has no direct physical significance. It will be clear, however, that at the temperature at which  $\theta'$  ascends so rapidly owing to the rapid increase in the time of relaxation, the  $\theta$  curve flattens out to its final value  $\theta_0$ . The temperature  $T_m$  at which  $\theta'$  has a minimum is therefore a useful index of the final temperature  $\theta_0$ , and this obviates the necessity of solving the differential equation (19) for each rate of cooling  $\sigma$ . We can, in fact, assume that  $\theta_0$  is equal to  $T_m$  without sensible error,\* where  $T_m$  is the temperature at which the curve  $\theta' = T + \sigma\tau$  has a minimum (see fig. 11). Hence

$$(d\theta'/dT)_{T=T_m} = 1 + \sigma (d\tau/dT)_{T=T_m} = 0.$$

Combining this with the relation  $\tau = Ae^{W/kT}$ ,

$$(AW/k\theta_0^2) e^{W/k\theta_0} = 1/\sigma.$$

Hence

$$W/k\theta_0 = \log(k\theta_0^2/\sigma AW),$$

\* In the actual curve plotted in the graph  $\theta_0$  is  $880^\circ$  K. and  $T_m$  is  $892^\circ$  K. This difference is negligible in the approximation aimed at in the present calculations, the more so because  $T_m$  is systematically somewhat higher than  $\theta_0$ .

or, since  $W/kT_1 = -\log A^{-1}$

$$\frac{\theta_0}{T_1} = \frac{1}{1 + \{\log (\theta_0^3 / \sigma T_1 \log A^{-1})\} / \log A^{-1}}. \quad (21)$$

Although  $\theta_0$  occurs on both sides of this equation, the fact that it occurs in a log term on the right-hand side makes it a simple matter to calculate  $\theta_0/T_1$ , for any rate of cooling  $\sigma$ . The expression has been put in this form to show that  $\theta_0$  is not very sensitive to changes in  $\sigma$ ,  $\log A^{-1}$  being considerably greater than the other logarithmic term. Values of  $\theta_0$  for different rates of cooling are given in Table III. A value of  $\sigma = 10^3$  corresponds to rapid quenching at the rate of  $1000^\circ$  per second. A value of  $\sigma = 10^{-3}$  is slow cooling at the rate of  $1^\circ$  every seventeen minutes. As before,  $A = 10^{-12}$  and  $T_1 = 1000^\circ \text{K}$ .

Table III.

$\sigma$ .	$\theta_0$ (absolute). ° K.	$\theta_0$ (Centigrade). ° C.
$10^3$	1125	852
$10^2$	1038	765
10	958	685
1	892	619
$10^{-1}$	833	560
$10^{-2}$	782	509
$10^{-3}$	737	464

The small difference made to these temperatures by assigning different values to  $A$  is illustrated by the following table.

Table IV.

	$A = 10^{-10}$ .	$A = 10^{-12}$ .	$A = 10^{-14}$ .
	° K.	° K.	° K.
$\sigma = 10^3$ ... ..	$\theta_0 = 1144$	1125	1110
$\sigma = 10^{-3}$ ... ..	695	737	768

If  $T_1$  has a higher or lower value than the  $1000^\circ \text{K}$ . assumed here, these estimates must be raised or lowered almost in proportion (the proportion would be exact if  $\sigma$  were also increased or decreased proportionately). They again illustrate what can be achieved by rapid quenching or very slow cooling. It is impossible to quench so rapidly as to preserve the alloy in a state corresponding

to  $\theta_0/T_1 > 1.2$ , or to anneal so slowly as to realize the state  $\theta_0/T_1 < 0.7$ . Intermediate states can be preserved by annealing for a time longer than the time of relaxation at the required temperature, and then quenching in a time considerably shorter than  $\tau$  (see Table II).

Kurnakow and Ageew's measurements of resistivity are of particular interest in view of these results. In fig. 10 the two experimental curves were obtained for different rates of cooling. The authors only give the total times of cooling as being 3 hours for the upper curve and 12–14 hours for the lower curve, but we may roughly estimate that they correspond to  $\sigma = 0.1$  and  $\sigma = 0.02$ . Judging by the small degree of order of the final state of the crystal, the corresponding  $\theta_0$  is not far below the transition temperature  $T_s$ , which is  $427^\circ \text{C}$ . ( $700^\circ \text{K}$ ). With these rates of cooling,  $\theta_0/T_1$  is about 0.8 (see Table III). Hence  $T_1$  for  $\text{AuCu}_3$  is about  $820^\circ \text{K}$ . or  $550^\circ \text{C}$ . It should, therefore, be easily possible to preserve the alloy in the disordered state by quenching, as is, in fact, found experimentally. Prolonged annealing, on the other hand, could hardly reduce the characteristic temperature  $\theta_0$  to much below  $580^\circ \text{K}$ . ( $0.7 T_1$ ), or  $273^\circ \text{C}$  at which the degree of order  $S$  has only risen to 0.75. In  $\text{AuCu}$  a rate of cooling  $\sigma = 0.02$  produces a degree of order 0.78, and this indicates a value for  $\theta_0$  of about  $350^\circ \text{C}$ ., and so a value for  $T_1$  of  $780^\circ \text{K}$ . or  $500^\circ \text{C}$ . This is still above the transition point  $T_s$  at  $460^\circ \text{C}$ ., and quenching preserves the disordered state. On the other hand, these estimates are in marked disagreement with an observation of the same authors, that an  $\text{AuCu}$  alloy in which the disordered structure had been preserved by rapid quenching, and which was subsequently heated, began to relax between  $150^\circ$  and  $200^\circ \text{C}$ .\* If  $T_1$  is  $500^\circ \text{C}$ ., relaxation should not be appreciable below  $270^\circ \text{C}$ . It is clear that the estimates can only be regarded as qualitative, but at the same time the picture they present is coherent.

The values of  $V$  for both structures, according to the simple theory, are somewhat over 0.3 electron-volts, and the values of the activation energy required for interchange of atomic position are 1.97 and 1.85 electron-volts for  $\text{AuCu}_3$  and  $\text{AuCu}$  respectively.

Haughton and Payne\* quote considerably lower values for the critical temperatures of both  $\text{AuCu}$  and  $\text{AuCu}_3$ . They obtain on cooling about  $400^\circ \text{C}$ . for the  $\text{AuCu}$  and  $380^\circ$  for  $\text{AuCu}_3$ . We have used the data of Kurnakow and Ageew, because they give the resistivity over the whole temperature range; we wish rather to indicate the lines along which the behaviour of an alloy can be explained than give figures for these special alloys.

\* 'J. Inst. Metals,' vol. 46, p. 475 (1931).

6. *Calculation of the Time of Relaxation.*

We require to find the rate at which the degree of order  $S$  approaches its equilibrium value  $S_e$ , when the actual degree of order is  $S_e + \delta S$ . The rate depends upon the difference between the number of  $\alpha\beta$  transitions and of  $\beta\alpha$  transitions per second (an  $\alpha\beta$  transition is one from an ordered position  $\alpha$  to a disordered position  $\beta$  and *vice versa*). We will only outline the calculation of  $\tau$ , in order to show how an estimate of the order of magnitude of the constants involved may be obtained.

The number,  $n_{\alpha\beta}$  of  $\alpha\beta$  transitions per second is proportional to the product of the number of replacements in  $\alpha$  positions, viz.,  $prn$ , and the probability that any given  $\beta$  position is not already occupied by a replacement, viz.,  $(1 - 2r + rp)/(1 - r)$ . Thus

$$n_{\alpha\beta} = f_{\alpha\beta} prn (1 - 2r + rp)/(1 - r).$$

The coefficient  $f_{\alpha\beta}$  is not strictly a constant, being indirectly dependent on  $p$  because the energy change during transition depends on the degree of order. The number of  $\beta\alpha$  transitions is similarly

$$n_{\beta\alpha} = f_{\beta\alpha} (1 - p) rn (1 - p).$$

At equilibrium  $n_{\alpha\beta} = n_{\beta\alpha}$

$$p (1 - 2r + rp)/(1 - r) (1 - p)^2 = f_{\beta\alpha}/f_{\alpha\beta}.$$

This way of considering the equilibrium closely resembles the application of the Law of Mass Action in chemical reactions,\* and  $f_{\beta\alpha}/f_{\alpha\beta}$  on the right-hand side may be regarded as the equilibrium constant,  $c$ . The latter obeys the thermodynamical relation

$$c = c' e^{q/kT},$$

where  $q$  is the decrease of potential energy for elementary reaction, and in our problem is the energy  $V$ . Thus

$$f_{\beta\alpha}/f_{\alpha\beta} = c' e^{V/kT}.$$

$c'$  depends upon special features of the process under consideration. In our problem it is determined by the fact that at  $T = \infty$  the distribution is random, i.e.,  $p = r$ . This gives  $c' = r/(1 - r)$ , so that

$$f_{\beta\alpha}/f_{\alpha\beta} = r/(1 - r) \cdot e^{V/kT}.$$

\* We may again emphasize that although we are using the Law of Mass Action as in chemical reactions, we are not dealing with the formation of molecules of a chemical compound. It is the relation of a given atom to the surrounding structure, not its relation to another atom, which is being considered.



It will be seen that with this value for  $f_{\beta\alpha}/f_{\alpha\beta}$  the equilibrium relation is identical with the relation (5) obtained in § 2.

Suppose now that the degree of order is slightly higher than that corresponding to equilibrium. Since the number of replacements in ordered positions is greater,  $n_{\alpha\beta}$  is greater than it is for equilibrium and similarly  $n_{\beta\alpha}$  is less. The number in ordered positions therefore drops back towards normal. Since  $p$  is the probability that a position of order is occupied by a replacement and  $nr$  is the total number of replacements as in section 3

$$\begin{aligned} nr \cdot dp/dt &= \delta n_{\beta\alpha} - \delta n_{\alpha\beta} \\ &= \frac{d}{dp} (n_{\beta\alpha} - n_{\alpha\beta}) \delta p, \end{aligned}$$

$p$  having increased from  $p_0$  to  $p_0 + \delta p$ .

By definition of the time of relaxation  $\tau$

$$dp/dt = -\delta p/\tau.$$

Hence

$$\tau^{-1} = -\frac{d}{dp} (n_{\beta\alpha} - n_{\alpha\beta})/nr.$$

Carrying out the differentiation and making the substitution  $p = r + S(1-r)$  leads to the equation

$$\tau^{-1} = f_{\alpha\beta} \frac{1 + S_0}{(1 - S_0)(1 - r)} \cdot \left(1 - \frac{m_s}{m_v}\right).$$

The factor  $(1 - m_s/m_v)$  can be understood by reference to fig. 2. If  $V$  remained a constant when  $S$  increases from  $S_0$  to  $S_0 + \delta S$ , the factor would be unity. Actually  $V$  increases and although with stable equilibrium the increase is not sufficient to maintain the higher degree of order it slows up the process of return, in the ratio  $(1 - m_s/m_v)$ , where  $m_s$  is the slope of the  $S(V, T)$  curve and  $m_v$  the slope of the  $V(S, T)$  curve. As we have already seen, if  $m_s > m_v$  the equilibrium is unstable. If  $m_s = m_v$  the time of relaxation becomes infinite because the new degree of order remains one of equilibrium. This has an interesting physical significance. It leads us to expect a lag in the response of the structure to a change in temperature near the critical temperature, where the angle of intersection of the  $V$  and  $S$  curves is very small. Such a lag is frequently observed (see Kurnakow and Ageew, *loc. cit.*) as a difference in the transition temperature for cooling and heating.

It is therefore seen that  $\tau^{-1}$  is equal to  $f_{\alpha\beta}$  multiplied by a numerical factor of the order unity. It remains to calculate  $f_{\alpha\beta}$ , which denotes the probability

per second that a replacement in a position of order  $\alpha$  makes a transition to a neighbouring position of disorder  $\beta$ . We shall suppose such a transition to be a simple interchange of position of two atoms A and B, which occurs when the following conditions are satisfied :—

- (1) That A passes through its equilibrium with kinetic energy greater than  $W/2$ , and in a direction which lies within a certain solid angle  $4\pi\omega$ .
- (2) That the same condition applies to B.
- (3) That these events agree in time within a fraction  $\phi$  of the period of oscillation, which may be taken to be the same for both.

In other words, an interchange will take place during a period of oscillation if both atoms have sufficient energy, are aimed correctly, and approach the barrier nearly simultaneously. These conditions lead to the following equation

$$f_{\alpha\beta} = \nu \omega^2 \phi q \cdot (8/3\pi) \cdot (W/kT)^2 \cdot e^{-W/kT}.$$

Here  $\nu$  is the number of oscillations of each atom per second, and  $q$  is the number of neighbouring places to which an A atom can make a transition. Strictly, both  $q$  and  $W$  depend upon the degree of order, but the factors which allow for this are so nearly unity that they can be neglected. Hence

$$\tau^{-1} = \frac{8}{3\pi} \cdot \frac{1 + S_e}{(1 - S_e)(1 - r)} \cdot \left(1 - \frac{m_i}{m_o}\right) \cdot \nu \omega^2 \phi q \left(\frac{W}{kT}\right)^2 e^{-W/kT}.$$

Since  $\nu$  is of such high order,  $W/kT$  must be large compared with unity in any case where  $\tau$  is appreciable. The variation of  $e^{-W/kT}$  with temperature is therefore much more important\* than any variation in the rest of the expression, and we may write

$$\tau = A e^{W/kT},$$

where  $A$  is effectively a constant.

To obtain the numerical value of  $A$  the following value of the factors are taken

$$\frac{8}{3\pi} \cdot \frac{1 + S_e}{(1 - S_e)(1 - r)} \cdot \left(1 - \frac{m_i}{m_o}\right) = 1.$$

It is of this order unless the last factor is very small

$\nu = 10^{12}$ , corresponding to a characteristic temperature of  $500^\circ \text{K}$ .

$$\omega^2 \phi = 5 \times 10^{-5}$$

$$q = 4$$

$$(W/kT)^2 = 600.$$

\* A parallel is found in the equation for thermionic emission.

This gives  $A = 10^{-12}$ . The value assigned to  $\omega^2\phi$  is based on  $\phi$  being  $1/20$ , and the assumption that the directions of motion of the two interchanging atoms as they pass through their positions of equilibrium  $a$  and  $b$  make an angle between  $30^\circ$  and  $45^\circ$  with the line  $ab$ , and also that their directions diverge on opposite sides of  $ab$  to within  $30^\circ$ . Such estimates merely suffice to give the order of  $A$ , but as we have seen this is all that is required since it is  $\log A$  which appears in the calculation of the preceding paragraphs. It would appear justifiable to estimate that  $A = 10^{-n}$  where  $n = 12 \pm 3$ . The order of  $A$  is, in fact, mainly determined by the frequency of vibration of the atoms, each atom being given  $10^{13}$  opportunities per second to surmount the barrier if direction and energy of motion are suitable. We have made similar calculations for a ring transition in which three atoms take part, and this leads to estimates of the same order of magnitude. In each system,  $W$  is the total energy involved in surmounting the barrier.

### *Conclusions.*

The distribution of the atoms amongst the sites in an alloy phase is the result of a dynamical equilibrium. Where a superlattice exists, the atoms of each kind are more often in certain positions of "order" than in alternative positions of "disorder." Any given atom, however, is continually wandering through the structure by interchanges between one phase-site and another. The regularity of the superlattice is statistical.

Although at room temperature the process of atomic interchange in most alloys has practically ceased, every alloy has in its previous history passed through a high temperature where atomic interchange is just sufficiently frequent to establish dynamic equilibrium. Its character has been impressed on it in this condition. When the temperature is lowered it fails to relax to its new equilibrium structure owing to the rapid decrease in the frequency of atomic interchange with fall of temperature. We must thus go back to the state of dynamic equilibrium at higher temperatures in order to explain the characteristics of the arrangement of atoms in the alloy at room temperature.

The peculiar form of the relation between the equilibrium degree of order and temperature is a general result of the interdependence of  $S$  and  $V$  discussed in this paper. To put it broadly, it is due to the assumption that  $S$  creates  $V$  and  $V$  creates  $S$ . Any such interaction leads to a critical temperature above which there is complete disorder, and below which there is an abrupt onset of partial order.

It has been assumed in the above discussion that the relative numbers of atoms of two kinds are those required for the ideal superlattice. It will be clear that the arguments advanced apply equally well if the ratios of the numbers of atoms are not exactly integral. This is in accord with observations on superlattice formation. The associated phenomena (change of resistivity, etc.) are most marked at the correct ratio, but appear in diminishing extent over a wide range of composition on either side.

Each superlattice has two important characteristics. They appear in the analysis as  $V_0$  (energy of interchange) and  $W$  (activation energy required for interchange). They are most simply represented, however, by two temperatures. The first is  $T_0$ , the critical temperature at which order appears. This governs the relation between temperature and the equilibrium degree of order. The second is  $T_1$ , the temperature at which the time of relaxation is 1 second. This determines the effects of quenching and annealing.

The possibility of obtaining an alloy in a state of disorder by quenching, and of order by annealing, depends upon the ratio between  $T_1$  to  $T_0$  being suitable.

The application of the terms "intermetallic compound" and "solid solution" to alloys such as Au-Cu is often debated. The point at issue would appear to be, not so much the physical nature of the relation between gold and copper in these alloys, if the point of view adopted here is accepted, but the justification for calling a relation of this kind a "compound." A description of alloys which depart from the correct ratio for AuCu, or  $\text{AuCu}_3$ , as "solid solutions of Au (or Cu) in the intermetallic compounds AuCu or  $\text{AuCu}_3$ " is definitely misleading, if it implies that some of the gold atoms, for example, are associated in a molecular way with copper and the remainder are in some way different ("in solution"). All atoms of each kind play an identical part in the dynamical equilibrium. The physical nature of the structure can be more adequately described by replacing the conceptions of "solid solution" and "intermetallic compound" by "phase-structure" and "atomic distribution."

We were led to develop this theory as the result of some observations on the Fe-Al alloys. A subsequent search in the literature of a branch of research quite unfamiliar to us showed that many previous experimental observations seemed to support the theory. At the same time we realized that in some respects we were following again lines of thought developed by previous workers, Tammann in particular. Nevertheless, we believe the present treatment to differ in several essential ways from other attempts to explain alloy structure

and in particular to lend itself to simple numerical calculations which give a satisfactory explanation of many experimental results.

[*Note added in proof, March 19, 1934.*—The rate of inter-diffusion of two metals, in which the atoms of one diffuse through the single crystal blocks of the other, is intimately connected with the time of relaxation of an alloy of the two metals. Actually the coefficient of diffusion,  $D$ ,  $\sim a^2/\tau$ , where  $a$  is the lattice spacing, and  $\tau$  is the time of relaxation. Where both  $D$  and  $\tau$  have been measured, we find that the coefficient of diffusion is of a different order of magnitude from  $a^2/\tau$ , being very much greater. We conclude either that the diffusion then takes place not inside single crystals, but between the single crystals in the metallic structure; or, as is less likely, that there occur atomic interchanges inside a crystal for which the potential barrier  $W$  is considerably less than that which applies to the atomic interchanges which alter the degree of order.]

We wish to thank Dr. H. Bethe and Dr. R. Peierls for helpful advice and criticism during the preparation of this paper.

### *Summary.*

An alloy phase has two characteristics. The first is the pattern of sites occupied by atoms irrespective of their nature. Each phase of an alloy system has a different pattern of sites, and therefore a change from one phase to another involves their complete re-arrangement. The second characteristic is the distribution of the atoms amongst these sites. This distribution may vary continuously without change of phase, from being random at high temperatures to being partially regular at low temperatures. The effect of thermal treatment upon the arrangement of the atoms forms the subject of the present paper.

The equilibrium states of the alloy are first considered, and the degree of order of the structure as a function of temperature is calculated. The ordered structure has a lower potential energy than the disordered structure, but thermal agitation promotes disorder. It is shown that above a certain critical temperature the structure is completely random. As the temperature is lowered, order sets in abruptly at the critical temperature, and at first rapidly increases. It only becomes complete as absolute zero is approached.

This characteristic sudden onset of order indicated by the theory causes a sharp inflexion in curves which show the variation of resistivity, lattice spacing

and specific heat with temperature. These inflexions simulate a phase-change, though there is actually no such change.  $\beta$  brass, Au-Cu alloys, and Fe-Al alloys are cited as examples.

The second section of the paper deals with the rate at which an alloy, not in equilibrium, relaxes towards equilibrium. A general law for the dependence of rate of relaxation upon temperature is deduced, which enables the effects of annealing and quenching to be predicted. The rate of relaxation depends upon the magnitude of the "activation energy" required to surmount a potential barrier when two atoms interchange position.

The alloy is a system of dynamical equilibrium. Although interchange of atomic position at room temperature is infrequent, the alloy has received its character at some previous point in its history when the temperature was just sufficiently high for interchange to be important. Maxima and minima in physical properties at certain relative proportions (*e.g.*, Fe<sub>3</sub>Al and AuCu<sub>3</sub>) are statistical effects, and do not imply the existence of corresponding compounds.

---

(Abstracts)

541.14 : 547.9  
535.61—31 : 621.38*Physico-Chemical Studies of Complex Organic Molecules. Part I.—  
Monochromatic Irradiation.*By F. P. BOWDEN and C. P. SNOW, Laboratory of Physical Chemistry,  
Cambridge.

(Communicated by T. M. Lowry, F.R.S. —Received December 27, 1933.)

A method is described for the production of monochromatic light of sufficient intensity to bring about reasonably rapid photochemical changes. The irradiation can be performed on very small amounts of material and the progress of the reaction followed spectroscopically. Selective monochromatic irradiation is applied to some of the large molecules of biological importance notably ergosterol and calciferol, vitamin B<sub>1</sub>, carotene and vitamin A.

(The full paper appears in 'Proc. Roy. Soc.,' B, vol. 115, p. 261 (1934)).

535.342 : 547.9

*Physico-Chemical Studies of Complex Organic Molecules. Part II.—  
Absorption Spectra at Low Temperatures.*By F. P. BOWDEN and S. D. D. MORRIS, Laboratory of Physical Chemistry,  
Cambridge.

(Communicated by T. M. Lowry, F.R.S. —Received December 27, 1933.)

The absorption spectra of some important biological molecules have been measured at liquid air temperature. The bands of  $\beta$  carotene (in ethyl alcohol) become narrower and shift to 4990 Å., 4670 Å., and 4350 Å., and a new band appears at 4060 Å. The ultra-violet band at 2700 Å. becomes sharper but is little displaced.

The main band of vitamin A concentrates at 3280 Å. is shifted to 3350 Å. and new bands appear at 2900 Å., 2770 Å., 2580 Å., 2510 Å., and 2430 Å.

The irradiation product of carotene whose absorption band at room temperature resembles that of vitamin A develops a structure at low temperatures with bands at 3780 Å., 3570 Å., 3410 Å., and 3210 Å. This more precise spectroscopic test shows that the two substances are not the same.

The absorption spectrum of vitamin E concentrates is due to several different molecules and some progress has been made in separating these out.

(The full paper appears in 'Proc. Roy. Soc.,' B, vol. 115, p. 274 (1934)).





## INDEX to VOL CXLV (A)

- Air, thermal conductivity (Hercus and Sutherland), 599
- Alloys, copper zinc, X ray examination (Owen and Pickup) 258
- Alloys, effect of thermal agitation on atomic arrangement (Bragg and Williams), 699
- Alloys, iron aluminium, peculiarities of physical properties (Sykes and Evans), 529
- Alloys, nickel copper, specific heat (Grew), 509
- Alpha particle groups, analysis of fine structure from thorium C (Lewis and Bowden), 235
- Alpha-rays, energies (Wilson), 447
- Aluminium chloride, band spectrum (Bhaduri and Fowler), 321
- Alumosilicate framework structures, nature and properties (Taylor), 80
- Argon, small angle inelastic electron scattering (Whiddington and Taylor) 465
- Atoms, inner shell ionization (Burhop), 612
- Bailey (C R) and Cassie (A B D) Investigations in the infra red region of the spectrum  
X—The asymmetrical molecule nitrosyl chloride NOCl, 336
- Bakerian lecture on the structure of alloys (Bragg), 699
- Bateson (S) See Henderson and Bateson
- Beilby layer (Finch and others), 676
- Bennett (G M) and Glasstone (S) An analysis of the dipole moments of some aromatic compounds, 71
- Beta rays, energies (Wilson), 447
- Bhaduri (B N) and Fowler (A) Band spectrum of aluminium chloride (AlCl<sub>3</sub>), 321
- Bitter (F) An elementary discussion of ferromagnetism, 629
- Bitter (F) On the fracture of fibred iron silicon sheets, 668
- Bowden (B V) See Lewis and Bowden
- Bowden (F P) and Morris (S D D) Physico chemical studies of complex organic molecules II—Absorption spectra at low temperature (Abstract) 731
- Bowden (F P) and Snow (C P) Physico chemical studies of complex organic molecules I—Monochromatic irradiation (Abstract) 731
- Bragg (W L) and Williams (E J) The effect of thermal agitation on atomic arrangement in alloys, 699
- Bulder (G) See Green and Bulder
- Burhop (E H S) The probability of inner shell ionization of atoms by electron impact, 612
- Cassie (A B D) See Bailey and Cassie
- Copper zinc alloys, X ray examination at elevated temperatures (Owen and Pickup), 258
- Cox (W M) and Wolfenden (J H) The viscosity of strong electrolytes measured by a differential method, 475
- Crystals, mechanism of plastic deformation (Taylor), 362, 388
- Denissoff (A K) and Richardson (O W) The emission of electrons under the influence of chemical action III—The action of Cl<sub>2</sub>, Br<sub>2</sub>, I<sub>2</sub>, NOCl, HCl, N<sub>2</sub>O, and COS on NaK<sub>2</sub>, 18

- Dielectric constants of phenol-water mixtures (Howell and Jackson), 539.
- Dielectrics, solid, theory of the electrical breakdown (Zener), 523.
- Dipole moments, analysis, of some aromatic compounds (Bennett and Glasstone), 71.
- Dobson (G. M. B.) *See* Götz and others.
- Dymond (E. G.) On the polarization of electrons by scattering, II, 657.
- Elasticity, volume, faults in a material which yields to shear stress while retaining (Taylor), 1.
- Electrolytes, strong, viscosity measured by a differential method (Cox and Wolfenden), 475.
- Electron impact, new transition produced in helium (Whiddington and Priestley), 462.
- Electron scattering, inelastic, in helium, neon and argon (Whiddington and Taylor), 465.
- Electron theory of metals (Schubin and Wonsowsky), 159.
- Electrons, emission under the influence of chemical action, III (Denisoff and Richardson), 18.
- Electrons, polarization by scattering (Dymond), 657.
- Evans (H.) *See* Sykes and Evans.
- Evans (R. C.) The atomic work function of tungsten for potassium, 135.
- Feachem (C. G. P.) *See also* Tronstad and Feachem.
- Feachem (C. G. P.) and Tronstad (L.) An optical examination of thin films. II.—The behaviour of thin films of fatty acids on mercury, 127.
- Ferromagnetism discussion (Bitter), 629.
- Field, magnetic, theory of change in resistance (Jones and Zener), 268.
- Films, metal, structure (Finch and Ikin), 551.
- Films, thin, optical examination (Tronstad and Feachem), 115, 127.
- Finch (G. I.) and Ikin (A. W.) The catalytic properties, and structure of metal films. II.—The electrical condition of platinum films, 551.
- Finch (G. I.), Quarrell (A. G.) and Roebuck (J. S.) The Beilby layer, 676.
- Flint (N. T.) A relativistic basis of the quantum theory, II, 645.
- Flow, measurement of turbulence (Townend), 180.
- Flow, turbulent, velocity variations (Simmons and Salter), 212.
- Fowler (A.) *See* Bhaduri and Fowler.
- Frameworks, tubular, having continuous longitudinals, stress calculation (Owen), 104.
- Gamma-rays, energies (Wilson), 447.
- Glasstone (S.) *See* Bennett and Glasstone.
- Götz (F. W. P.), Meetham (A. R.) and Dobson (G. M. B.) The vertical distribution of ozone in the atmosphere, 416.
- Green (A. L.) and Builder (G.) The rotation of the plane of polarization of long radio waves, 145.
- Grew (K. E.) The specific heat of nickel and of some nickel-copper alloys, 509.
- Haloes, pleochroic (Henderson and Bateson and Henderson and Turnbull), 563, 582.
- Helium, new transition produced by electron impact (Whiddington and Priestley), 462.
- Helium, small angle inelastic electron scattering (Whiddington and Taylor), 465.
- Henderson (G. H.) A new method of determining the age of certain minerals, 591.
- Henderson (G. H.) and Bateson (S.) A quantitative study of pleochroic haloes, I, 563.
- Henderson (G. H.) and Turnbull (L. G.) A quantitative study of pleochroic haloes, II, 582.
- Hercus (E. O.) and Sutherland (D. M.) The thermal conductivity of air by a parallel plate method, 599.

- Hooker (S. G.) Two-dimensional oscillations in divergent jets of compressible fluid, 52.
- Howell (O. R.) and Jackson (W.) The dielectric constants of liquid mixtures of phenol—water, phenol—*m*-cresol, phenol—*o*-aniline, and phenol—*p*-toluidine, 539.
- Ikin (A. W.) *See* Finch and Ikin.
- Iron-aluminium alloys, peculiarities of physical properties (Sykes and Evans), 529.
- Iron-silicon sheets, fracture of fibred (Bitter), 668.
- Jackson (W.) *See* Howell and Jackson.
- Jones (H.) and Zener (C.) The theory of the change in resistance in a magnetic field, 268.
- Lewis (W. B.) and Bowden (B. V.) An analysis of the fine structure of the  $\alpha$ -particle groups from thorium C and of the long range groups from thorium C', 235.
- Meetham (A. R.) *See* Götz and others.
- Mercury, behaviour of thin films of fatty acids on (Feachem and Tronstad), 127.
- Mercury, optical constants (Tronstad and Feachem), 115.
- Metals, electron theory (Schubin and Wonsowsky), 159.
- Methane and oxygen, reaction when sensitized by nitrogen peroxide (Norrish and Wallace), 307.
- Methylammonium cyanate, rate of transformation in aqueous solution into methylurea (Miller), 288.
- Methylurea, rate of transformation in aqueous solution of methylammonium into (Miller), 288.
- Miller (C. C.) The rate of transformation in aqueous solution of methylammonium cyanate into methylurea, 288.
- Minerals, determination of age (Henderson), 591.
- Molecules, physico-chemical studies of complex organic (Bowden and Snow), and (Bowden and Morris), 731.
- Murti (S. G. K.) *See* Rao and Murti.
- Neon, small angle inelastic electron scattering (Whiddington and Taylor), 465.
- Neutron, quantum theory (Temple), 344.
- Nickel-copper alloys, specific heat (Grew), 509.
- Nitrogen dioxide, structure of molecule from study of its infra-red absorption spectrum (Sutherland), 278.
- Nitrogen peroxide, reaction of methane and oxygen sensitized by (Norrish and Wallace), 307.
- Nitrosyl chloride, the asymmetrical molecule NOCl (Bailey and Cassie), 336.
- Norrish (R. G. W.) and Wallace (J.) The reaction of methane and oxygen sensitized by nitrogen peroxide. I.—Thermal ignition, 307.
- Oscillations, two-dimensional, in divergent jets of compressible fluid (Hooker), 52.
- Owen (E. A.) and Pickup (L.) X-ray examination of certain copper-zinc alloys at elevated temperatures, 258.
- Owen (J. B. B.) Stress calculations for tubular frameworks having continuous longitudinal, 104.
- Oxygen, reaction when sensitized by nitrogen peroxide (Norrish and Wallace), 307.
- Ozone, vertical distribution in atmosphere (Götz and others), 416.
- Phenol—water, phenol—*m*-cresol, phenol—*o*-aniline and phenol—*p*-toluidine mixtures, dielectric constants (Howell and Jackson), 539.

- Pickup (L.) *See* Owen and Pickup.
- Polarization plane of long radio waves, rotation (Green and Builder), 145.
- Potassium, atomic work function of tungsten (Evans), 135.
- Priestley (H.) *See* Whiddington and Priestley.
- Quantum theory of the neutron (Temple), 344.
- Quantum theory, relativistic basis (Flint), 645.
- Quarrell (A. G.) *See* Finch, Quarrell, and Roebuck.
- Radio waves, long, rotation of plane of polarization (Green and Builder), 145.
- Rao (I. Ramakrishna) The constitution of water in different states, 489.
- Rao (K. R.) and Murti (S. G. K.) Investigations on the spectrum of selenium. III.—Extension of Se III; IV.—Se I and Se VII, 681, 694.
- Richardson (O. W.) *See* Denisoff and Richardson.
- Roebuck (J. S.) *See* Finch, Quarrell and Roebuck.
- Salt, rock, strength (Taylor), 405.
- Salter (C.) *See* Simmons and Salter.
- Schubin (S.) and Wonsowsky (S.) The electron theory of metals, 159.
- Selenium spectrum of Se III, and Se I and Se VII (Rao and Murti), 681, 694.
- Simmons (L. F. G.) and Salter (C.) Experimental investigation and analysis of the velocity variations in turbulent flow, 212.
- Spectrum, band, of aluminium chloride (AlCl<sub>3</sub>), 321.
- Spectrum, investigations in the infra-red region (Bailey and Cassie), 336.
- Spectrum of selenium (Rao and Murti), 681, 694.
- Stabilizer, field, high sensitivity automatic magnetic (Wynn-Williams), 250.
- Stress calculation for tubular frameworks having continuous longitudinals (Owen), 104.
- Stress, shear, faults in a material which yields to (Taylor), 1.
- Sutherland (D. M.) *See* Hercus and Sutherland.
- Sutherland (G. B. B. M.) Structure of the molecule of nitrogen dioxide from a study of its infra-red absorption spectrum, 278.
- Sykes (C.) and Evans (H.) Some peculiarities in the physical properties of iron-aluminium alloys, 529.
- Taylor (G. I.) Faults in a material which yields to shear stress while retaining its volume elasticity, 1.
- Taylor (G. I.) The mechanism of plastic deformation of crystals, I, II, 362, 388.
- Taylor (G. I.) The strength of rock salt, 405.
- Taylor (J. E.) *See* Whiddington and Taylor.
- Taylor (W. H.) The nature and properties of aluminosilicate framework structures, 80.
- Temple (G.) The quantum theory of the neutron, 344.
- Thorium C and C', analysis of the fine structure of the  $\alpha$ -particle groups (Lewis and Bowden), 235.
- Townend (H. C. H.) Statistical measurements of turbulence in the flow of air through a pipe, 180.
- Tronstad (L.) *See also* Feachem and Tronstad.
- Tronstad (L.) and Feachem (C. G. P.) An optical examination of thin films. I.—The optical constants of mercury, 115.
- Tungsten, atomic work function for potassium (Evans), 135.
- Turbulent flow, experimental investigation and analysis of the velocity variations (Simmons and Salter), 212.

Turbulence, statistical measurements in the flow of air through a pipe (Townend), 180.  
 Turnbull (L. G.) *See* Henderson and Turnbull.

Velocity variations in turbulent flow, investigation and analysis (Simmons and Salter), 212.  
 Viscosity of strong electrolytes (Cox and Wolfenden), 475.

Wallace (J.) *See* Norrish and Wallace.

Water, constitution in different states (Rao), 489.

Whiddington (R.) and Priestley (H.) Note on a new transition produced by electron impact in helium, 462.

Whiddington (R.) and Taylor (J. E.) Small angle inelastic electron scattering in helium neon, and argon, 465.

Williams (E. J.) *See* Bragg and Williams.

Wilson (H. A.) The energies of alpha, beta, and gamma rays, 447.

Wolfenden (J. H.) *See* Cox and Wolfenden.

Wonsowsky (S.) *See* Schubin and Wonsowsky.

Wynn-Williams (C. E.) An automatic magnetic field stabilizer of high sensitivity, 250.

Zener (C.) A theory of the electrical breakdown of solid dielectrics, 523.

Zener (C.) *See* Jones and Zener.

END OF THE ONE HUNDRED AND FORTY-FIFTH VOLUME (SERIES A)

---



IMPERIAL AGRICULTURAL RESEARCH  
INSTITUTE LIBRARY  
NEW DELHI.

Date of issue.	Date of issue.	Date of issue.
6-11-57		
10-7-57		
JUL 1954		
NOV 1954		
2-61		
-12-61		
7-61		